



HAL
open science

A la recherche d'anticorps humains et de nouveaux biomarqueurs pour le ciblage de la plaque d'athérome

Audrey Hemadou

► **To cite this version:**

Audrey Hemadou. A la recherche d'anticorps humains et de nouveaux biomarqueurs pour le ciblage de la plaque d'athérome. Immunologie. Université de Bordeaux, 2017. Français. NNT : 2017BORD0806 . tel-01714482

HAL Id: tel-01714482

<https://theses.hal.science/tel-01714482>

Submitted on 21 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE

Pour le

Doctorat de l'Université de Bordeaux
Mention : Science de la Vie et de la Santé
Option : Immunologie, Microbiologie

TITRE

A la recherche d'anticorps humains et de nouveaux biomarqueurs pour le ciblage de la plaque d'athérome

Présentée et soutenue publiquement le
05/12/2017

par

HEMADOU Audrey

Composition du jury :

Pr. Zahi FAYAD, Professeur hospitalo-universitaire	Rapporteur
Dr. Brigitte KERFELEC, Chargée de recherche CNRS	Rapporteur
Pr. Eric DUCASSE, Chef de service « chirurgie vasculaire »	Examineur
Pr. Vincent DOUSSET, Professeur hospitalo-universitaire	Président
Dr. Véronique GIUDICELLI, Ingénieur de recherche CNRS	Examineur
Dr. Alexandre FONTAYNE, Responsable d'unité de recherche LFB	Examineur
Dr. Gisèle CLOFENT-SANCHEZ, Directrice de recherche CNRS	Examineur
Dr. Marie-Josée JACOBIN-VALAT, Ingénieur de recherche CNRS	Examineur

Thèse préparée au sein de l'UMR5536 CNRS – Université de Bordeaux

« Tout obstacle renforce la détermination. Celui qui s'est fixé un but n'en change pas. »

Léonard de Vinci, Scientifique, Ingénieur, Les Carnets, 1987

Remerciements

Voici le point final à une recherche de plus de trois années qui sans un soutien incontestable de ma famille n'aurait jamais pu être écrit.

Dans un premier temps mes remerciements vont à mes deux rapporteurs, le Pr Zahi FAYAD ainsi que le Dr. Brigitte KERFELEC d'avoir pris le temps de juger ce travail malgré les impératifs de chacun. Je remercie également le reste des membres du jury pour leur intérêt vis-à-vis de ces travaux.

Je tiens à remercier tout particulièrement le Pr Zahi FAYAD, non seulement pour les discussions autour d'une collaboration entre nos deux équipes mais également de m'avoir interviewée pour un éventuel post-doctorat dans son équipe au Mont Sinaï aux Etats-Unis.

Je souhaite remercier Alexandre FONTAYNE et Martine CERUTTI pour leurs précieux conseils prodigués durant les nombreuses réunions ANR autour du projet ATHERANOS qui a financé pour moitié ma thèse. Merci de m'avoir aiguillée et soutenue dans mes recherches, d'avoir pris le temps de m'écouter lors de mes présentations orales et d'avoir validé chacun de mes résultats. Je remercie également Véronique GIUDICELLI ainsi que Marie-Paule LEFRANC pour leur aide autour du projet de séquençage haut débit en passant par la rédaction de l'article scientifique jusqu'à la soumission des séquences NGS et à la préparation des fichiers. Merci au Pr DUCASSE qui, par son grand intérêt vis-à-vis du projet, a contribué à ma réussite ainsi qu'à la proposition de perspectives fortes intéressantes de l'utilisation de ces anticorps en clinique.

Bien sûr, après avoir remercié les membres du jury ainsi que les collaborateurs, la première personne à qui je dédie ce travail, c'est toi, Marie-Josée JACOBIN-VALAT pour ta patience, ton encadrement et ta sagesse scientifique qui m'ont guidés vers le chemin de la réussite. Ta bonne humeur au quotidien et ta positivité ont été d'une grande aide dans la vie de tous les jours au laboratoire ainsi que dans les moments difficiles. Tu es et resteras la directrice de thèse qu'il me fallait, ce que j'ai su dès l'instant où je t'ai rencontré lors de mon choix de stage de master 2. Bien plus qu'une directrice de thèse, tu as été ma confidente pendant ces trois années. Je savais que je pouvais parler du travail mais également de sujets tout aussi divers avec toi pour nous changer les idées.

Un grand merci à Mme Gisèle CLOFENT-SANCHEZ de m'avoir accueillie à bras ouvert dans son équipe et de m'avoir encadrée notamment sur le projet de séquençage haut débit avec Pacific Biosciences ainsi qu'à Mme Jeanny LAROCHE-TRAINEAU pour ces heures de manipulations passer à ses côtés à apprendre son savoir-faire sur de nombreuses techniques de biologie moléculaire. Je tiens à vous remercier toutes les deux, ainsi que Florence OTTONES et Cyril LORENATO pour vos conseils et corrections pour ce manuscrit de thèse.

Un grand merci à Ségolène ANTOINE pour sa rigueur expérimentale et un soutien inconditionnel au cours de ces trois années. Cette rencontre au sein de cette unité a été une révélation autant sur le plan professionnel que personnel. Toujours là dans les bons comme les mauvais moments, tu resteras et je resterais toujours à tes côtés. J'ai trouvé ma confidente et une sœur de cœur « ma bichette ».

Merci au Pr Jean-Michel FRANCONI et Mr Sylvain MIRAUX de m'avoir accueillie au sein du laboratoire ainsi que Mr Pierre DOS SANTOS de nous avoir alloués les locaux de la Plateforme Technologique d'Innovations Biomédicales. Je souhaiterais également remercier Mr Alain GADEAU et tous ceux qui ont participé à la réussite de ces trois années de recherche.

J'aimerais ensuite exprimée ma gratitude envers ma famille à qui je dédie également ce travail, à ma maman, pour son soutien sans faille, mon frère et ma grand-mère qui m'ont permis d'arriver jusque-là ainsi que mon oncle pour ces bonnes paroles. Je voudrais remercier Josiane DANIELOU et Jacky DANIELOU, qui a rejoint les étoiles, d'avoir été présents sur Bordeaux quand ma famille était loin de moi. A mon papa et mes deux grands-pères, j'ai toujours cru en votre soutien de là-haut.

Enfin mes pensées vont vers toi, mon Amour, qui m'a soutenu, réconforté et apporté des conseils durant ces trois années. Tu as toujours su me supporter malgré mon humeur massacrant de temps à autre et mon sale caractère. Merci à toi... A nous l'Australie. On en rêvait... On y est ...

Table des matières

Liste des abréviations.....	9
Liste des figures.....	12
Liste des tableaux.....	14
Avant-propos.....	15
Preface.....	16
Chapitre 1: L'athérogénèse.....	18
I Les artères.....	18
I.1 Généralités.....	18
I.2 Les constituants de la paroi saine.....	18
I.3 Différents types d'artère.....	19
I.3.1 Les artères élastiques.....	19
I.3.2 Les artères musculaires.....	19
I.3.3 Les artérioles.....	19
II L'athérome.....	19
II.1 Athérogénèse.....	19
II.2 Incidence des maladies cardiovasculaires à l'échelle mondiale.....	20
II.3 La classification des plaques d'athérome.....	21
II.4 Les conséquences de l'athérosclérose.....	22
II.4.1 Infarctus du myocarde.....	22
II.4.2 Accident vasculaire cérébral (AVC).....	22
II.4.3 Angine de poitrine (Angor).....	23
II.4.4 Anévrisme.....	23
III Acteurs cellulaires et facteurs moléculaires impliqués dans la formation de la plaque d'athérome.....	23
III.1 Acteurs cellulaires.....	23
III.1.1 Acteurs cellulaires non immuns.....	23
III.1.1.i Cellules endothéliales vasculaires.....	23
III.1.1.ii Cellules musculaires lisses (CML).....	23
III.1.1.iii La matrice extracellulaire.....	24
III.1.2 Acteurs cellulaires immuns.....	25
III.1.2.i Monocytes et macrophages.....	25
III.1.2.ii Les lymphocytes T.....	26
III.1.2.iii Les lymphocytes B.....	28
III.1.2.iv Les plaquettes.....	29
III.2 Facteurs moléculaires.....	29
III.2.1 Le cholestérol.....	29
III.2.2 Les espèces réactives oxygénées.....	32
III.2.3 Les cytokines.....	32
III.2.4 Les métalloprotéinases de la matrice (MMP).....	34
III.3 Conclusions.....	34
Bibliographie.....	35
Chapitre 2 : Vers l'imagerie de la plaque d'athérome.....	43
I Imagerie invasive de la plaque.....	43
I.1 L'angiographie.....	43
I.2 La thermographie.....	43

I.3	L'échographie intravasculaire (IVUS).....	44
I.4	La tomographie par cohérence optique (OCT).....	44
II	Imagerie non invasive de la plaque.....	45
II.1	La spectroscopie proche infra-rouge (NIRS).....	45
II.2	La fluorescence proche infra-rouge (NIRF).....	46
II.3	La tomodensitométrie (CT = Computed tomography).....	46
II.1	Tomographie par émission de positons (TEP).....	47
II.2	Tomographie par émission monophotonique (SPECT).....	47
II.3	Imagerie par résonance magnétique (IRM).....	48
III	Conclusion	52
	Bibliographie.....	55
	Chapitre 3: Les anticorps	60
I	Histoire.....	60
II	Les anticorps	61
II.1	Structure d'une immunoglobuline.....	61
II.2	Les domaines constants	62
II.3	Les domaines variables.....	62
II.4	Production d'anticorps par les lymphocytes B.....	62
III	La diversité des anticorps.....	65
III.1	Les familles multigéniques codant pour les immunoglobulines et leur réarrangement.....	66
III.2	La flexibilité jonctionnelle.....	68
III.2.1	La P addition	68
III.2.2	La N addition.....	68
III.3	L'hypermutation somatique.....	69
III.4	Association aléatoire des chaînes lourdes et légères	70
III.5	Conclusions	70
IV	Les différents formats d'anticorps	70
V	Les différents systèmes de production d'anticorps.....	71
V.1	La glycosylation des anticorps	72
V.2	Le système mammifère.....	73
V.3	Le système bactérien	74
V.4	Le système levure et champignon filamenteux	74
V.5	Le système des cellules d'insectes.....	74
V.6	Les plantes transgéniques	75
V.7	Les animaux transgéniques.....	75
VI	Les anticorps en thérapie	75
VI.1	Les anticorps murins (-momab).....	76
VI.2	Les anticorps chimériques (-ximab)	77
VI.3	Les anticorps humanisés (-zumab)	77
VI.4	Les anticorps recombinants humains (-mumab ou -umab)	77
VI.5	Applications thérapeutiques des anticorps	78
VII	Sélection d'anticorps par la technologie du « phage display ».....	79
VII.1	Introduction sur le phage filamenteux M13.....	79
VII.2	Cycle d'amplification du phage M13	80
VII.3	Exposition des protéines de fusion à la surface du phage.....	81
VII.4	Les phagemides.....	81
VII.5	Les banques combinatoires d'anticorps.....	83

VII.5.1	Construction d'une banque combinatoire.....	83
VII.5.2	Les banques synthétiques et semi-synthétiques.....	83
VII.5.3	Les banques naïves.....	84
VII.5.4	Les banques immunes.....	84
VII.6	Phage display <i>in vitro</i>	85
VII.6.1	Sélection sur un antigène isolé.....	85
VII.6.2	Sélection sur cellules.....	86
VII.6.3	Sélection sur sérum de patients.....	87
VII.6.4	Conclusions.....	87
VII.7	Phage display <i>in vivo</i>	88
VIII	Séquençage des anticorps.....	89
	Bibliographie.....	93
	Chapitre 4: Résultats.....	103
	Objectifs de la thèse.....	103
	1 ^{ère} partie : Criblage <i>in vitro</i> d'anticorps.....	104
I	Phage display <i>in vivo</i>	104
I.1	Introduction.....	104
I.2	Sélection <i>in vivo</i> de fragments d'anticorps.....	104
I.2.1	Le modèle animal utilisé : le lapin athéromateux.....	105
I.2.2	La sélection <i>in vivo</i> réalisée.....	105
I.2.3	Production de phages recombinants.....	106
I.2.4	Test de réactivité.....	106
	I.2.4.i Cytométrie de flux.....	106
	I.2.4.ii Immunohistochimie.....	108
I.3	Résultats et discussion.....	108
II	Découverte de la cible.....	167
II.1	Introduction.....	167
II.2	Validation <i>in vitro</i> d'anticorps humains après reformatage sous format scFv-Fc.....	167
II.2.1	Test de réactivité.....	167
	II.2.1.i La cytométrie de flux.....	167
	II.2.1.ii Immunohistochimie sur coupes de tissus athéromateux.....	168
	II.2.1.iii Immunoprécipitation suivie d'une analyse en spectrométrie de masse.....	168
II.3	Reformatage d'anticorps en vue d'un couplage aux nanoparticules sous format scFv-Fc-2C.....	169
	II.3.1.i Immunofluorescence sur coupes de tissus humains.....	170
	II.3.1.ii Macrofluorescence sur aortes de souris ApoE KO.....	170
II.4	Résultats et discussion.....	170
	2ème partie: Criblage <i>in silico</i> d'anticorps humains.....	234
I	Séquençage haut débit de phage-scFv issus d'une sélection <i>in vivo</i>	234
I.1	Introduction.....	234
I.2	Séquençage d'anticorps issus d'une sélection par phage display <i>in vivo</i>	235
I.2.1	Amplification par PCR de clones scFv.....	235
I.2.2	Préparation et séquençage des scFv amplifiés.....	237
I.3	Résultats et discussion.....	238
	Bibliographie.....	280

Sélection d'anticorps humains et découverte de nouveaux biomarqueurs	286
Couplage des anticorps humains à des nanoparticules pour une imagerie non-invasive de la plaque.....	289
Bibliographie.....	291

Liste des abréviations

ABCA1 : *ATP-cassette subfamily A1*
ABCG1 : *ATP-cassette subfamily G1*
ACAT1 : *acétyl-CoA-acetyltransférase*
Acm : *Anticorps monoclonaux*
ADCC : *Antibody dependent cell-mediated cytotoxicity*
ADCP : *Antibody dependent cellular phagocytosis*
ADM : *Adrenomédulline*
ADP : *Adénosine diphosphate*
AGE : *Advanced glycation end-products*
AHA : *American Heart Association*
AMM : *Autorisation de Mise sur le Marché*
AP-1 : *Activator protein 1*
ApoA1 : *Apolipoprotéine A1*
ApoB : *Apolipoprotéine B*
ApoE : *Apolipoprotéine E*
AVC : *Accident vasculaire cérébral*
BAFF : *B-cell activating factor*
BCR : *B-cell receptor*
BLyS : *B lymphocyte stimulator*
BmSPX : *Brugia malayi SXP protein*
CAV1 : *Cavéoline 1*
CCS : *Circular consensus sequence*
CDC : *Complement dependent cytotoxicity*
CDR : *Complementary Determining Region*
CES1 : *Carboxylestérase 1*
CETP : *Protéine de transfert d'ester de cholesteryl*
CHO : *Chinese hamster ovarian*
CHU : *Centre hospitalier universitaire*
CLIO : *CrossLinked IronOxide*
CML : *Cellules musculaires lisses*
CT : *Computed tomography*
CTSD : *Cathepsine D*
DC : *Dendritic cells*
DOTA : *Tétra-azacyclododécane-tétra-acétate*
DTPA : *Diéthylène-triamine-penta-acétate*
EGFR : *Epidermal Growth Factor Receptor*
ELISA : *Enzyme linked immuno assay*
EMA : *European medicines agency*
F2 : *Prothrombine*
FDA : *Food drug agency*
F-FDG : *F-fluorodesoxyglucose*
FR : *Framework*
GAG : *Glycosaminoglycanes*
GMCSF : *Granulocyte Macrophage Colony Stimulating Factor*
GSEA : *Gene set enrichment analysis*
HAMA : *Human anti mouse antibody*
HDL : *Lipoprotéines de forte densité*
HEK : *Human embryonic kidney*
HER2 : *Human Epidermal Growth Factor Receptor-2*
HTRF : *Homogeneous Time Resolved Fluorescence*
ICAM1 : *Intracellular Adhesion Molecule 1*
IFN γ : *Interferon gamma*

IL : *Interleukin*
 IMGT : *ImmunoGeneTics, ImmunoGeneTics*
 IP-10 : *Inteferon gamma induced protein - 10*
 IRA : *Innate Response Activator*
 IRM : *Imagerie par résonance magnétique*
 ITAC : *Interferon-inducible T-cell Alpha Chemoattractant*
 ITGB2 : *Intégrine β 2*
 IVUS : *Intravascular Utra-Sound*
 JNK : *c-Jun N-terminal kinases*
 KO : *Knockout*
 LCAT : *Lécithine cholestérol acétyltransférase*
 LDL : *Lipoprotéines de faible densité*
 LDLR : *Récepteurs aux LDLs*
 LFA1 : *Lymphocyte function-associated antigen*
 LPS : *Lipopolysaccharide*
 LRP1 : *Protéine reliée aux récepteurs de LDL*
 LXR : *Liver X receptor*
 MAGP : *Microfibrillar-associated glycoprotein*
 MCP1 : *Monocyte chemoattractant protein – 1*
 MCSF : *Macrophage Colony Stimulating Factor*
 MEC : *Matrice extracellulaire*
 MEPPS : *Morphology Enhanced Probability maPS*
 Mig : *Monokine induced by gamma interferon*
 MMP : *Matrix metalloproteinases*
 NADPH : *Nicotinamide adénine dinucléotide phosphate*
 NF- κ B : *Nuclear factor – kappa B*
 NGS : *Next generation sequencing*
 NIRF : *Fluorescence proche infra-rouge*
 NIRS : *Spectroscopie proche infra-rouge, Voir*
 NK : *Natural killer*
 NO : *Monoxyde d'azote*
 NS1 : *NonStructural protein 1*
 OCT : *Tomographie par cohérence optique*
 PCR : *Polymerase Chain Reaction*
 PCSK9 : *Proprotéine convertase subtilisine/kexine de type 9*
 PDGF : *Platelet Derived Growth Factor*
 PEG : *Polyéthylène glycol*
 PFA-4 : *Platelet Factor-4*
 PGI2 : *Prostacycline*
 Pim-1 : *Proto-oncogene serine/threonine-protein kinase*
 PON3 : *Paroxonase 3*
 PPAR : *Peroxisome proliferator-activated receptor*
 PSGL1 : *P-selectin glycoprotein ligand 1*
 RAG : *Recombination-Activating Gene*
 RANTES : *Regulated upon Activation, Normal T-cell Expressed and Secreted*
 rHDL : *Lipoprotéine de haute densité reconstituée*
 ROS : *Reactive oxygen species*
 RSS : *Séquences signal de recombinaison*
 scFv : *Single chain fragment variable*
 SGS : *Second Generation Sequencing*
 SLN : *Solid Lipid Nanoparticles*
 SMRT : *Single Molecule Real Time*
 SPECT : *Tomographie par émission monophotonique*

SPIO : *SuperParamagnetic IronOxide*
S-RA : *A scavenger receptor*
TCEP : *Tris(2-carboxyethyl)phosphine*
TCFA : *Thin-Cap Fibroatheroma*
TCR : *T Cell Receptor*
TdT : *Terminal déoxynucléotidyl*
TEP : *Tomographie par émission de positons*
TGFβ : *Transforming Growth Factor bêta*
TGS : *Third Generation Sequencing*
TOF : *Temps de vol*
TRAP-6 : *Peptide-6 activateur de récepteur de thrombine*
TXA2 : *Thromboxane A2*
UFM1 : *Ubiquitine 1*
USPIO : *Ultrasmall superparamagnetic iron oxide*
VGLUT : *Vesicular Glutamate Transporter*
VH-IVUS : *Vascular Histology – IVUS*
VLA4 : *Very late antigen 4*
VLDL : *Lipoprotéines de très faible densité*
VUSPIO : *Versatile UltraSmall SuperParamagnetic IronOxide*
vWF : *Von Willebrand Factor*
YAC : *Yeast Artificial Chromosome*

Liste des figures

Figure 1 Trois couches de cellules constituent la paroi des artères	18
Figure 2 Développement de l'athérosclérose.	20
Figure 3 Représentation schématique des différents stades nommés par l'AHA.	22
Figure 4 Mécanismes régulant le recrutement des monocytes et leur accumulation dans la plaque d'athérome.....	26
Figure 5 Les effets des sous-populations de cellules T sur l'athérosclérose.	27
Figure 6 Réponses des lymphocytes B dans l'athérosclérose.	28
Figure 7 Adhésion plaquette-endothélium.....	29
Figure 8 Le métabolisme des lipoprotéines joue un rôle clé dans l'athérogénèse	31
Figure 9 Rôle des ROS dans le développement de l'athérosclérose.	32
Figure 10 Coronarographie : pose d'un stent.....	43
Figure 11 Images IVUS délimitant la lumière d'un vaisseau sain et d'un vaisseau pathologique.....	44
Figure 12 Imagerie OCT des différents aspects de la plaque.	45
Figure 13 Image co-enregistrée d'IVUS en échelle de gris et NIRS obtenue grâce à un cathéter combiné IVUS-NIRS.....	46
Figure 14 Images de tomographie par cohérence optique (OCT) couplée à la fluorescence proche infrarouge (NIRF) de l'artère de lapin athéroscléreuse <i>in vivo</i>	46
Figure 15 Imagerie de la plaque carotidienne par la technique microPET avec du ¹⁸ F-fluorodeoxyglucose (FDG) <i>ex vivo</i>	47
Figure 16 Sections transversale (gauche), coronale (milieu) et sagittale (droite) de l'image micro-SPECT d'un échantillon de plaque carotidienne athéroscléreuse après une incubation avec ^{99m} Tc-folate montrant une accumulation hétérogène.	48
Figure 17 Principe de l'IRM.....	49
Figure 18 Caractérisation structurale des plaques d'athérosclérose par une combinaison de séquences.	50
Figure 19 Imagerie IRM d'une aorte abdominale de souris ApoE KO après injection de micelles fonctionnalisées à l'aide d'anticorps anti-LDLox	51
Figure 20 Imagerie IRM d'une aorte abdominale de souris ApoE KO après injection de rHDL chargées en statines et gadolinium.....	52
Figure 21 Les cibles d'imagerie moléculaire potentielles dans l'athérosclérose.....	54
Figure 22 Représentation schématique de la structure d'une immunoglobuline.....	61
Figure 23 Les domaines variables VH et VL des chaînes lourdes et légères d'une Ig ou d'un anticorps.....	62
Figure 24 Production d'anticorps à partir des cellules B matures.	64
Figure 25 Synthèse d'une chaîne lourde mu (μ) et d'une chaîne légère kappa d'immunoglobuline.	65
Figure 27 Représentation graphique des séquences RS.....	67
Figure 26 Représentation schématique de la formation de l'épingle à cheveux lors de la recombinaison V-J pour une chaîne légère.	67
Figure 28 Schématisation de la P et N addition.....	69
Figure 29 Différents formats d'anticorps.....	71
Figure 30 Tableau des anticorps présents actuellement sur le marché	72
Figure 31 Avantages comparatifs des systèmes de production d'anticorps monoclonaux	72
Figure 32 Profil de glycosylation des anticorps naturels.	73
Figure 33 Structure de l'anticorps humain et représentation des distinctions entre chaque type d'anticorps: Souris, Chimérique, Humanisé et Humain.....	76
Figure 34 La technique des hybridomes.....	76
Figure 35 Souris transgéniques produisant des immunoglobulines humaines, exemple des Xenomouse®.	78
Figure 36 La structure du phage M13 et de son génome.....	80
Figure 37 Assemblage et exportation de phages M13	81

Figure 38 Vecteur phagemide.....	82
Figure 39 Le concept de l'hyperphage: schématisation du phage auxiliaire déplété du gène de P3 dit « hyperphage » suivi de l'infection classique d'une bactérie.....	83
Figure 40 Phage display <i>in vitro</i> sur un antigène isolé et immobilisé.....	85
Figure 41 Phage display <i>in vitro</i> sur un antigène biotinylé adsorbé sur des billes/streptavidine.....	86
Figure 42 Phage display sur des cellules exprimant le récepteur d'intérêt.....	87
Figure 43 Phage display <i>in vivo</i> chez l'Homme.....	89
Figure 44 Comparaison de deux méthodes de travail pour la découverte d'anticorps.....	91
Figure 45 Tableau comparatif des trois générations de séquençage.....	92
Figure 46 Résumé et comparaison des nouvelles méthodes de séquençage haut débit utilisées à l'heure actuelle.....	92
Figure 47 Sélection par phage display <i>in vivo</i> d'une banque d'anticorps humains.....	106
Figure 48 Principe de fonctionnement de la cytométrie de flux.....	107
Figure 49 Schématisation des conditions (A à D) testées en cytométrie de flux pour la mise au point du criblage de phage-scFv.....	108
Figure 50 Sous-réseaux de processus cellulaires enrichis dans des échantillons d'aorte de lapins NZW athéroscléreux comparés à des lapins témoins NZW.....	111
Figure 51 Principe de la méthode d'immunoprécipitation.....	169
Figure 52 Exemple schématique d'un spectromètre de masse.....	169
Figure 53 Schématisation du principe d'immunofluorescence.....	170
Figure 54 Annexe 1 : test par cytométrie de flux des anticorps reformatés sous format scFv-Fc.....	172
Figure 55 Annexe 2 : immunohistochimie sur coupes de lapin athéromateux.....	173
Figure 56 Schématisation de l'amplification PCR effectuée avec placement des amorces sur le vecteur.....	235
Figure 57 Résultats PCR des essais de différentes amorces, nombres de cycles et dilutions.....	236
Figure 58 Résultat de l'amplification PCR des 4 produits avec l'utilisation de la polymérase Kapa Hifi Hot Start et 15 cycles.....	236
Figure 59 Résultat du Bioanalyzer sur les 4 produits PCR envoyés à Pacific Biosciences.....	237
Figure 60 Schématisation de la préparation des brins d'ADN et processus de séquençage.....	238
Figure 61 Processus de sélection d'anticorps humains par deux méthodes <i>in silico</i> et <i>in vitro</i> pour aboutir à la découverte de nouveaux biomarqueurs et à une meilleure compréhension de l'athérosclérose.....	288
Figure 62 Inhibition de l'agrégation plaquettaire par des SLN chargées en PGI2.....	290

Liste des tableaux

Tableau 1. Rôle pro- ou anti- athérogène des chimiokines.....	33
Tableau 2. Rôle pro- ou anti- inflammatoire des interleukines	33
Tableau 3 Nombre de gène V, D et J pour les chaînes légères κ et λ et les chaînes lourdes	66
Tableau 4. Tableau récapitulatif de l'enrichissement des phage-scFvs au cours des différents tours de sélection (P1 – P3) dans les différentes fractions.....	109
Tableau 5. Quantité de phage-scFv élués après les différents tours de sélection (P1 à P3) dans les différentes fractions	235

Avant-propos

Ce projet de recherche s'inscrit dans le cadre de la thématique « ciblage de la plaque d'athérome », portée par le Dr Gisèle Clofent-Sanchez au sein de l'UMR5536. Cette thématique est au carrefour de plusieurs disciplines et fait appel à des compétences multiples (physique, chimique et biologie) pour le développement de stratégies innovantes pour le diagnostic et la thérapie, soit le théranostic de l'athérosclérose. Le travail que je décris ici est donc un travail d'équipe au cours duquel j'ai appris par l'expertise de chacun à accroître mes connaissances sur la sélection d'anticorps humains ciblant la plaque d'athérome et la découverte de nouveaux biomarqueurs, pour *in fine*, les coupler à des nanoparticules pour améliorer le diagnostic et la prise en charge des patients à risque de rupture.

Au cours des quelques chapitres décrits ci-dessous que vous me faites l'honneur de juger, j'ai voulu faire un tour d'horizon sur les connaissances accumulées ces dernières années sur l'athérogénèse. Dans un premier temps, j'ai voulu exposer le processus d'athérogénèse qui implique un grand nombre de types cellulaires immuns et non immuns ainsi que différents facteurs moléculaires. Grâce à l'évolution de la connaissance de la plaque d'athérome, nous savons aujourd'hui que, plus que par l'aspect sténosant de la plaque d'athérome, la composition moléculaire et cellulaire joue un rôle déterminant dans la rupture de cette dernière. Dans un second temps, j'ai voulu présenter les différentes techniques d'imagerie d'aujourd'hui et de demain. De nos jours, il n'existe aucune technique d'imagerie non invasive (faisant appel à des nano-objets ciblant) permettant d'accéder à la composition cellulaire et moléculaire caractérisant la vulnérabilité de la plaque. C'est dans ce cadre que la recherche internationale s'est mobilisée pour le ciblage et la découverte de nouveaux biomarqueurs exprimés au sein des plaques d'athérome pour une meilleure prise en charge des patients à risque. Pour ce faire, l'imagerie non invasive, à laquelle j'ai consacré un chapitre, est en plein essor et notamment avec le développement d'objets fonctionnalisés par des anticorps humains. L'avènement de la biologie moléculaire ainsi que les connaissances autour de la technologie des anticorps ont permis l'élaboration de méthodologies ultra-performantes telles que l'obtention d'anticorps humains spécifiques par phage display. L'utilisation d'anticorps humains est d'autant plus souhaitable que les anticorps murins induisent fréquemment une immunogénicité. De plus les anticorps humains offrent la possibilité d'être directement transférables des études précliniques aux études cliniques.

Après avoir abordé ces différents aspects, j'exposerai le travail auquel je me suis consacrée durant ces trois années de thèse.

La première partie fait d'état d'une nouvelle technologie de **cytométrie de flux à moyen débit** pour le criblage ***in vitro* d'anticorps humains** préalablement isolés par **phage display *in vivo***. Par la suite, la détermination de la **cible** a été réalisée par **immunoprécipitation et spectrométrie de masse**.

La deuxième partie concerne une nouvelle technique de criblage *in silico* directement par séquençage haut débit des anticorps humains sélectionnés par phage display *in vivo*.

Le but final étant ici de définir des **couples anticorps/cible à visée théranostique, c'est-à-dire utiles à la fois pour le diagnostic et la thérapie d'une pathologie d'intérêt**.

Cette thèse a donné lieu à trois articles, deux soumis dans des journaux à comité de lecture internationale (*Frontiers in Immunology* et *Scientific Reports*) et un en cours de soumission. De plus, l'un des anticorps découvert par ces deux techniques fait l'objet d'un brevet déposé en octobre 2017.

Preface

This research project is part of the research thematic "target atheroma plaque", directed by Dr. Gisèle Clofent-Sanchez in UMR5536. This theme is at the crossroads of several disciplines and uses multiple skills (physics, chemistry and biology) for the development of innovative strategies for the diagnosis and the therapy, named the theranostic of atherosclerosis. The work I describe here is therefore a team effort in which I learned, through the expertise of each, to increase my knowledge on the selection of human antibodies targeting atheroma plaque and the discovery of new biomarkers, to ultimately, couple them to nanoparticles to improve the diagnosis and the managing of high risk patients.

In the few chapters described below, you do me the honor to judging, I wanted to provide an overview on the knowledge accumulated in recent years in atherogenesis. At first, I wanted to expose the process of atherogenesis that involves a large number of immune and non-immune cell types as well as different molecular factors. Thanks to the evolution in the knowledge of atheroma plaque, today we know that, more than the narrowing of the artery, the molecular and cellular composition plays a determining role in the rupture of the plaque. Afterwards, I wanted to present imaging procedures of today and tomorrow. Nowadays, non-invasive imaging methods using targeting nanoparticles able to decipher the molecular and cellular composition underlying the vulnerability of these plaques do not exist. Within this framework, the discovery and the targeting of new biomarkers of atherosclerosis has become a priority of international research for better management of patients at risk. To reach this aim, non-invasive imaging, to which I have devoted a chapter, is in full development and in particular with the development of objects functionalized with human antibodies. The advent in molecular biology and the knowledge on antibody technology have led to the elaboration of ultra-performing methodologies such as obtaining specific human antibodies by phage display. The use of human antibodies is highly desirable as murine antibodies frequently induce immunogenicity. In addition, human antibodies offer the possibility of being directly transferable from preclinical to clinical studies.

Having discussed these different topics, I will describe the work I have devoted myself during these three years of my thesis.

The first part reports the development of a new method of a **high throughput flow cytometry** for the ***in vitro* screening of human antibodies** previously obtained after ***in vivo* phage display bio-panning**. The identification of targets was accessed by **immunoprecipitation and mass spectrometry**.

The second part is about a new method of ***in silico* screening by Next-Generation Sequencing (NGS) of human antibodies issuing from the *in vivo* phage display selection**.

The ultimate goal is to define antibody / target pairs for the theranostic, that means both the diagnosis and the therapy of a pathology of interest.

This thesis resulted in three articles, two submitted in international peer-reviewed journals (*Frontiers in Immunology* and *Scientific Reports*) and one in the process of submission. In addition, one of the antibodies discovered by these two techniques is being protected by a patent filed in October 2017.

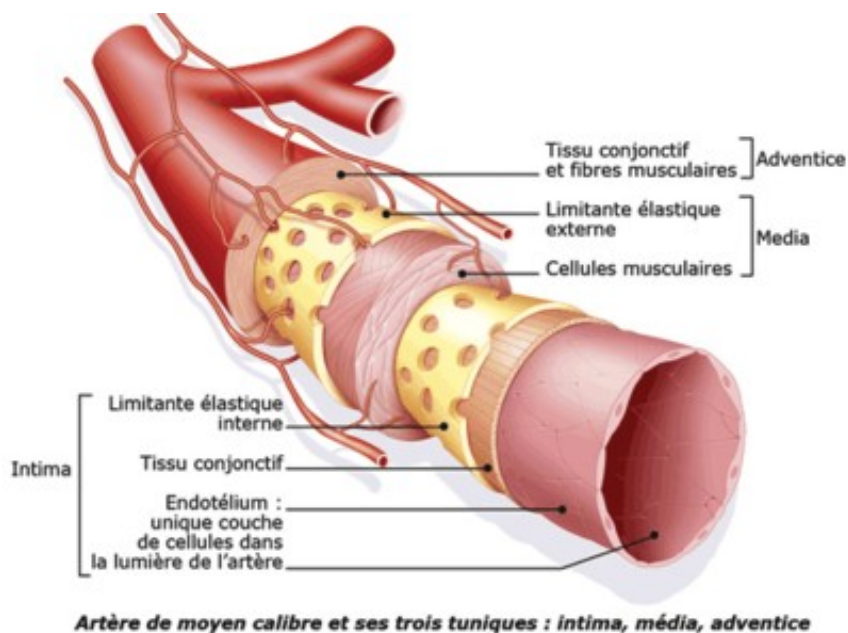
Introduction générale

Chapitre 1: L'athérogénèse

I Les artères

I.1 Généralités

Les artères, ou vaisseaux, ont pour fonction d'approvisionner les organes périphériques en sang provenant du cœur. Seules les artères de gros et moyen calibre sont exposées au risque de formation de plaques d'athérome. De plus, certaines zones sont préférentiellement concernées notamment les zones de courbures où le flux est réorienté et où la paroi est soumise à des contraintes mécaniques différentes. En termes d'anatomie, les artères sont constituées de trois couches différentes de cellules : de l'intérieur vers l'extérieur, l'intima, la média puis l'adventice (**Figure 1**).



Artère de moyen calibre et ses trois tuniques : intima, média, adventice

Figure 1 Trois couches de cellules constituent la paroi des artères : l'intima, la média et l'adventice (Extrait du site internet de Sanofi-Aventis, France)

I.2 Les constituants de la paroi saine

L'intima, partie la plus interne de l'artère, est constituée d'une monocouche de cellules endothéliales connectées entre elles par des jonctions serrées la rendant ainsi imperméable à toute molécule circulant dans le sang. Dans la partie sous endothéliale, la lame basale est principalement constituée de tissu conjonctif composé de fibroblastes et de composants de la matrice cellulaire. La limitante élastique interne définit la fin de l'intima et le début de la média. Cette dernière est composée de cellules musculaires lisses (CML) qui synthétisent des fibres élastiques et des fibres de collagène. Dans le cadre d'un tissu aortique, la média est constituée de lames concentriques dans lesquelles on retrouve les CML. Suivant l'artère considérée, elle peut parfois être séparée de l'adventice par une limitante élastique externe. S'ensuit donc l'adventice, riche en fibres de collagènes, en fibres élastiques épaisses et en fibroblastes. Le vaso vasorum, ayant pour rôle l'irrigation de la partie externe de la média, se situe dans cette couche. Deux dispositions de fibres élastiques sont observées : parallèlement à l'axe du vaisseau pour renforcer la structure en formant des trousseaux de fibres et perpendiculairement à l'axe du vaisseau de manière à ancrer le vaisseau au tissu avoisinant.

I.3 Différents types d'artère

Il existe trois différents types d'artère : les artères élastiques, les artères musculaires et les artérioles ayant des fonctions bien distinctes.

I.3.1 Les artères élastiques

Les artères élastiques correspondent aux artères de gros calibres (aorte, artères pulmonaires, carotides et sous-clavières). Ces artères sont principalement composées de fibres élastiques afin de pouvoir supporter les fortes augmentations de pression dues à l'éjection systolique du cœur (éjection, après contraction du cœur, du sang dans la circulation systémique). La paroi est peu épaisse si on prend en considération le diamètre du vaisseau.

I.3.2 Les artères musculaires

Leur fonction principale est de distribuer le sang aux différents organes et d'adapter le débit sanguin à la région concernée (exemple des coronaires). Il s'agit donc ici d'avoir des propriétés contractiles importantes avec modification du calibre ce qui se traduit par la présence d'une média plus épaisse riche en cellules musculaires lisses. La paroi est ici plus épaisse par rapport au diamètre du vaisseau.

I.3.3 Les artérioles

La structure des artérioles est la structure la plus simplifiée. L'intima, la média et l'adventice sont constituées chacune d'une seule couche de cellules. Elles ont pour rôle de transporter le sang vers les capillaires. Les artérioles sont dépourvues de vaso vasorum.

II L'athérome

L'athérosclérose, maladie chronique et inflammatoire des artères de gros et moyen calibre, se caractérise par l'accumulation de lipides (triglycérides et cholestérol) et de débris fibreux formant une plaque au sein de l'intima.

II.1 Athérogénèse

Depuis la plus jeune enfance jusqu'à l'adolescence, des lésions intimes non pathologiques causées par le stress hémodynamique se développent en l'absence de lipides. Les cellules musculaires lisses (CML) viennent alors épaissir la paroi des artères en particulier dans les zones de courbures plus susceptibles au changement directionnel du flux sanguin¹. A un âge plus avancé, sous l'influence de facteurs extérieurs tels que le tabagisme, l'obésité ou l'hypercholestérolémie, des lipoprotéines de faible densité (LDL) viennent former des stries lipidiques au sein de la paroi endothéliale endommagée. S'ensuit alors une réaction immunitaire dans ces zones d'entrée des LDL où elles vont être oxydées (**Figure 2**). Les cellules endothéliales activées vont exprimer des molécules d'adhésion telles que la P-sélectine qui va permettre le recrutement de monocytes dans la région sub-endothéliale. Dans ce microenvironnement, les plaquettes vont également participer au recrutement de leucocytes (granulocytes, monocytes et lymphocytes) et de LDL oxydées (LDLox)^{1,2,3}. Les plaquettes expriment de nombreuses molécules comme la P-sélectine quand elles sont activées mais également des molécules d'adhésion telles que ICAM-1 (Intracellular Adhesion Molecule 1) ou des glycoprotéines comme les intégrines $\alpha\text{IIb}\beta\text{3}$ ($\text{GP}\alpha\text{IIb}\beta\text{3}$) ou $\text{Ib}\alpha$ ($\text{GPIb}\alpha$)². Un état progressif de l'athérogénèse s'installe dû non seulement à une surcharge de LDL dans le sang mais également à une réponse exacerbée du système immunitaire localement. Les LDL s'accumulent dans les CML, dans les cellules endothéliales et dans la matrice extracellulaire (MEC) où elles subissent une oxydation par les ROS (reactive oxygen species) générés dans

cet environnement inflammatoire. Lors du processus d'oxydation, l'apolipoprotéine B (ApoB) est oxydée ce qui va conduire à sa reconnaissance par les macrophages résidents. Cette reconnaissance entraîne une phagocytose de ces LDL par les macrophages qui forment alors des cellules spumeuses ^{4,5,6}. La sécrétion par les CML et les macrophages de molécules chimioattractantes comme MCP-1 (Monocyte chemoattractant protein – 1) permet le recrutement de lymphocytes T, B mais aussi de monocytes, de neutrophiles et de cellules dendritiques dans la paroi artérielle. L'accumulation de cholestérol notamment dans les cellules spumeuses va générer la formation d'un corps nécrotique dans lequel un état d'hypoxie (manque d'oxygène) se met en place. Les macrophages vont entrer en apoptose et sécréter des molécules telles que les métalloprotéinases nommées « MMP (matrix metalloproteinases) » ou des collagénases qui vont fragiliser la paroi pouvant entraîner sa rupture. La fissure de la plaque expose les constituants de la MEC qui ont un rôle thrombogène. Un thrombus se forme et va évoluer localement en ischémie ou alors, migrer soit vers le cœur et causer un infarctus du myocarde, soit vers le cerveau et causer un accident vasculaire cérébral (AVC) ^{7,8}. De manière alternative, lors de la coagulation, la thrombine et le PDGF (Platelet Derived Growth Factor) générés entraînent le recrutement des CML tandis que le TGFβ (Transforming Growth Factor bêta), sécrété par les plaquettes activées, stimule la production de collagène. La prolifération des CML épaissit la chape fibreuse ce qui engendre un rétrécissement de la lumière du vaisseau ⁹.

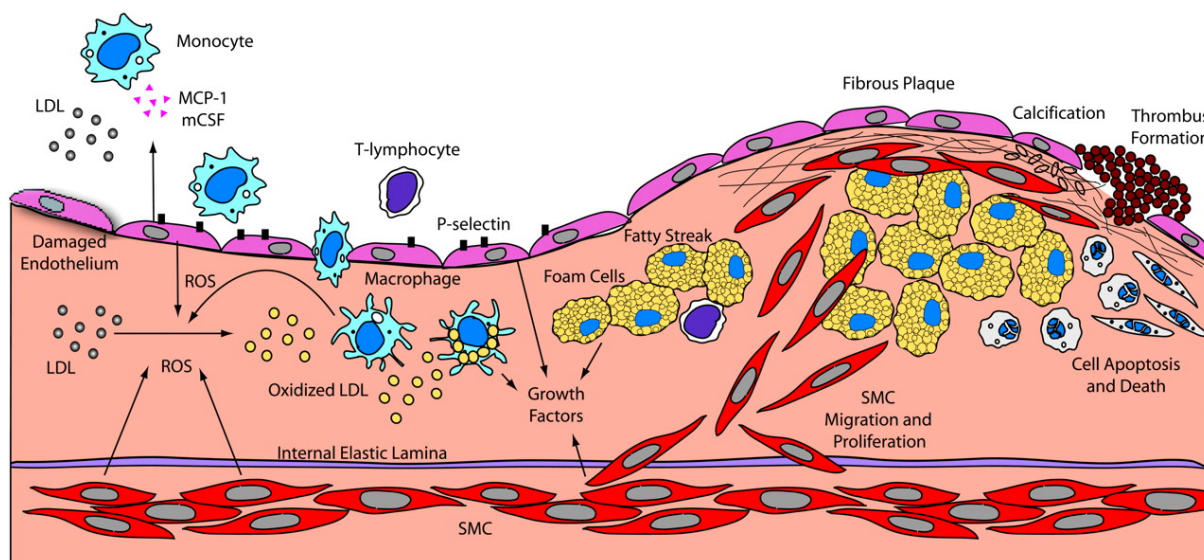


Figure 2 Développement de l'athérosclérose. Les ROS produits par les cellules endothéliales, les CML et les macrophages oxydent les LDL dans l'espace sous-endothélial, au niveau de l'endothélium endommagé, déclenchant des événements qui vont aboutir à la formation d'une plaque fibreuse. La rupture de la plaque fibreuse conduit à la formation de thrombus et à l'occlusion du vaisseau (Extrait de l'article de Madamanchi et al. paru en 2005 dans *Arterioscle Thromb Vasc Biol.*)

II.2 Incidence des maladies cardiovasculaires à l'échelle mondiale

Selon l'OMS (http://www.who.int/cardiovascular_diseases/en/, chiffres de 2012 consultés le 10/2017), il est rapporté que les maladies cardiovasculaires dont le facteur principal est l'athérosclérose sont la première cause de décès dans le monde. Environ 17 millions de personnes décèdent des suites de maladies cardiovasculaires, soit 31% de la mortalité totale mondiale. Au cœur de cette mortalité, les deux principales causes sont les coronaropathies avec 7,4 millions de morts et les accidents vasculaires cérébraux avec 6,4 millions de morts. Plusieurs facteurs interviennent notamment le tabagisme, l'hypertension artérielle, l'hypercholestérolémie, l'obésité et l'alcoolisme. Les maladies cardiovasculaires se regroupent dans plusieurs catégories :

- Les coronaropathies (touchant les vaisseaux qui alimentent le muscle cardiaque)

- Les maladies cérébro-vasculaires telles que les AVC (touchant les vaisseaux qui alimentent le cerveau)
- Les artériopathies périphériques (touchant les vaisseaux qui alimentent les bras et jambes telles que les ischémies)
- Les cardiopathies rhumatismales, affectant le muscle et les valves cardiaques et résultant d'un rhumatisme articulaire aigu, causé par une bactérie streptocoque
- Les malformations cardiaques congénitales (malformations de la structure du cœur déjà présentes à la naissance)
- Les phlébites (obstruction des veines des jambes par un caillot sanguin, susceptible de migrer vers les poumons (embolies pulmonaires)).

II.3 La classification des plaques d'athérome

La classification des plaques d'athérome en 6 stades plus ou moins évolutifs a été établie par l'American Heart Association (AHA) en 1994-1995 (**Figure 3**)¹⁰. Les lésions de type I et de type II s'observent dès l'enfance et consistent en un remaniement de la paroi artérielle notamment au niveau de l'intima. Le stade I correspond à une lésion initiale avec un épaissement de l'intima au niveau des zones de courbures, bifurcations ou branches artérielles. Elle se caractérise par quelques amas de macrophages spumeux au sein de la paroi endothéliale. La lésion de type II commence lors de l'apparition de la strie lipidique qui correspond à une accumulation de macrophages spumeux. Les cellules musculaires migrent à travers la limitante élastique interne. Il y a apparition de dépôts lipidiques parsemés. La lésion de type III se caractérise par la présence de dépôts lipidiques extracellulaires dans l'épaissement intimal et est appelée lésion pré-athéromateuse. Au stade III, les lésions peuvent encore être réparées et une régression est encore possible. La lésion de type IV est une lésion athéromateuse. Les caractéristiques typiques de la plaque apparaissent avec la présence d'un cœur lipidique (accumulation et confluence de dépôts lipidiques) composé de cholestérol libre ou estérifié, de phospholipides, des cellules spumeuses, de débris cellulaires nécrotiques et de micro-thrombi. Autour de ce cœur nécrotique se forme un anneau scléreux constitué d'une chape fibreuse avec une sécrétion importante de matrice extracellulaire et la présence de cellules musculaires lisses. L'évolution va alors conduire à la lésion de type V. Il s'agit alors d'un fibroathérome. Trois sous-types se distinguent : le stade Va correspond au stade IV avec une chape fibreuse plus épaisse composée de cellules musculaires lisses et de fibres de collagène ; le stade Vb voit l'apparition de dépôts calcifiés ; le stade Vc est caractérisé par une prédominance de la fibrose. La dernière évolution de la plaque est la lésion de type VI. De même que le stade V, le stade VI est composé de sous-types : un stade d'ulcération avec une érosion de la plaque et fissure de la chape nommé VIa. La plaque est considérée comme vulnérable, instable à ce stade. Lorsqu'il y a hémorragie à partir du vasa vasorum, la lésion est dite au stade VIb. Le dernier stade VIc est caractérisé par le thrombus mettant en jeu les plaquettes. Le thrombus peut causer une embolie, une occlusion ou alors se stabiliser en s'intégrant à la plaque en créant un rétrécissement de la lumière du vaisseau plus importante.

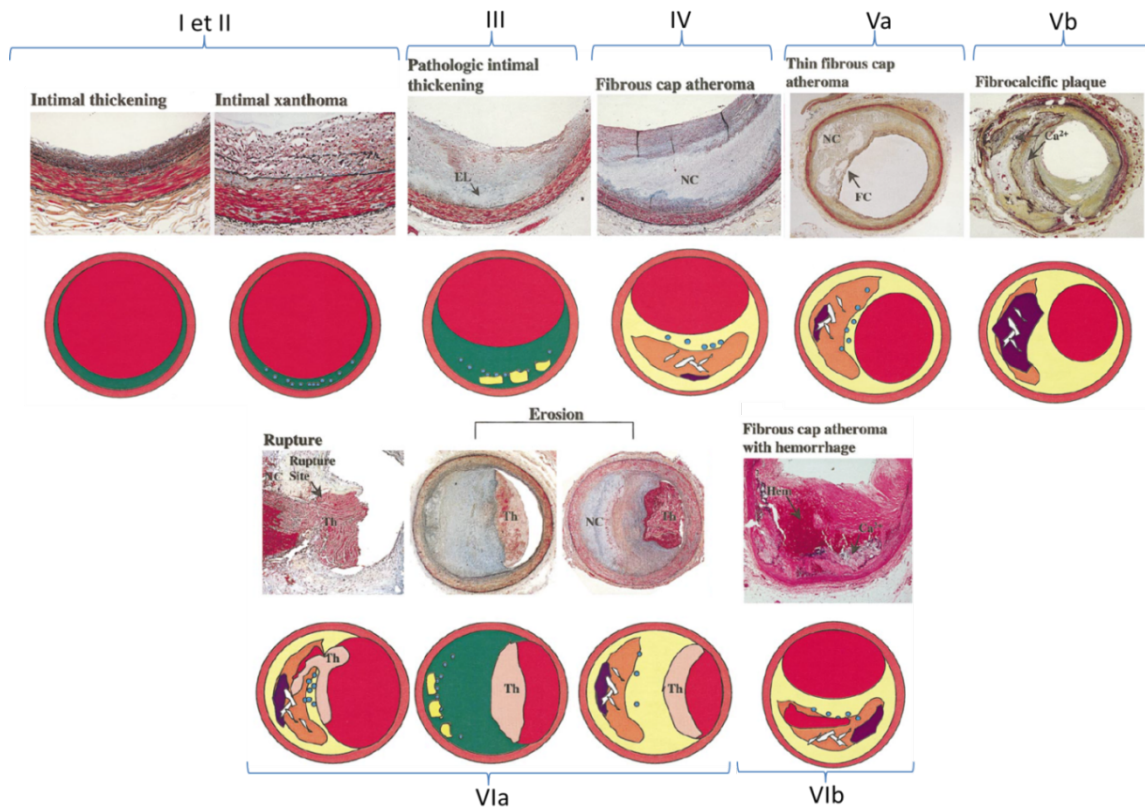


Figure 3 Représentation schématique des différents stades nommés par l'AHA (Figure modifiée de l'article de Virmani et al. paru en 2000 dans *Arterioscle Thromb Vasc Biol.*).

II.4 Les conséquences de l'athérosclérose

II.4.1 Infarctus du myocarde

L'infarctus du myocarde survient après l'obstruction d'une artère qui alimente le cœur en sang et donc en oxygène. En état d'anoxie (absence d'oxygène), les cellules meurent et rentrent en apoptose excessive causant la nécrose d'une zone déterminée. Suite à cette nécrose, le muscle cardiaque présente des problèmes de contraction ce qui entraîne une arythmie cardiaque suivie d'une insuffisance cardiaque pouvant mener à l'arrêt du cœur. L'infarctus du myocarde se traduit par une oppression prolongée de la poitrine, douleur aiguë dans le thorax qui peut irradier dans le bras, la mâchoire, la gorge et l'épaule. La douleur est traitée par des dérivés nitrés ou des antalgiques opiacés.

II.4.2 Accident vasculaire cérébral (AVC)

L'accident vasculaire cérébral est lié à un manque d'apport en oxygène dans une zone du cerveau. Il peut être dû à une thrombose (occlusion) de la carotide interne ou à une embolie cérébrale (migration d'un caillot ou d'un débris gras provenant d'une rupture de plaque d'athérome). Il se traduit par une paralysie (déficit neurologique). La paralysie peut concerner tout l'hémicorps (hémiplégie) ou un membre (inférieur ou supérieur) avec en association une paralysie faciale ou trouble du langage. Suivant le temps qui s'écoule entre la formation de l'anévrisme et la prise en charge du patient, plusieurs traitements peuvent être mis en place, allant de la dissolution du thrombus (thrombolyse) à l'endartériectomie carotidienne.

II.4.3 Angine de poitrine (Angor)

L'angine de poitrine se manifeste lors d'un effort. Elle est due à un rétrécissement du vaisseau subissant un remodelage de la paroi interne. Ceci se traduit par une diminution de l'apport de sang et donc d'oxygène. Le cœur a donc du mal à effectuer sa fonction principale. De manière à réduire la douleur, la trinitrine, agissant sur le relâchement des CML, est utilisée.

II.4.4 Anévrisme

Il a été remarqué, à proximité des anévrismes, la présence de plaques d'athérome. Ainsi, anévrisme et plaques d'athérome pourraient être liés. L'anévrisme est dû à une malformation vasculaire cérébrale acquise, correspondant à une dilatation focale d'une artère cérébrale. L'anévrisme peut se rompre et induire une hémorragie interne qui peut être fatale. Dans la majorité des cas l'anévrisme, s'il est silencieux, fait l'objet d'une surveillance ou peut être également opéré chirurgicalement de manière à prévenir sa rupture chez les personnes à risque (fort taux de cholestérol, hypertension, hyperlipidémie...).

III Acteurs cellulaires et facteurs moléculaires impliqués dans la formation de la plaque d'athérome

III.1 Acteurs cellulaires

III.1.1 Acteurs cellulaires non immuns

III.1.1.i Cellules endothéliales vasculaires

La fonction primaire de l'endothélium est de maintenir le sang dans les artères. Il est décrit comme la première barrière, le premier filtre moléculaire permettant l'échange de substances nutritives avec le milieu intérieur. Il est impliqué également dans la régulation du fonctionnement des tissus sous-jacents. L'endothélium participe à la relaxation des CML par synthèse de monoxyde d'azote (NO). L'endothélium est directement soumis aux forces de cisaillements générées par le fluide en écoulement permanent. Ainsi, ces forces de cisaillements agissent sur l'orientation et la forme des cellules endothéliales. Des expériences d'écoulement sur un lit de cellules endothéliales en culture ont montré que l'activité métabolique et le taux de prolifération de ces cellules sont modifiés¹¹. Ces contraintes pourraient modifier le flux de calcium pénétrant dans la cellule de la paroi vasculaire et provoquer une libération de substances vasoactives. Lors d'une contraction, le taux de calcium augmente grâce aux vasoconstricteurs tels que l'angiotensine ou l'endothéline. Lors d'une relaxation, le processus inverse se produit. Le taux de calcium baisse par l'action des vasodilatateurs tels que l'adrénaline ou bien l'ADM (Adrenomédulline)¹². Une autre fonction rattachée à l'endothélium est le contrôle de la coagulation sanguine par synthèse de dérivés de l'acide arachidonique, comme la prostacycline (PGI₂)¹³. L'endothélium joue un rôle déterminant dans l'inflammation car il exprime des molécules d'adhésion (telles que les sélectines P et E) et des molécules de la superfamille des immunoglobulines (ICAM-1 et 2 et VCAM-1) permettant le recrutement des leucocytes.

III.1.1.ii Cellules musculaires lisses (CML)

Les cellules musculaires lisses correspondent à des cellules allongées, fusiformes et mononucléées. Leur cytoplasme comporte des protéines contractiles, actine et myosine, permettant de contrôler la contraction. Les CML s'organisent en faisceaux. Elles sont liées entre elles par des desmosomes et des jonctions maculaires *macula adherens*. Les CML

produisent des constituants de la matrice extracellulaire tels que le collagène, les fibres élastiques ou les protéoglycanes. La contraction des cellules musculaires lisses est dépendante du calcium et de l'ATP. Toute modification de la concentration intracellulaire en calcium peut entraîner des contractions plus ou moins fréquentes des CML avec des conséquences sur la pression artérielle. Une pression artérielle trop basse conduira à un apport en O₂ faible suite à un ralentissement du débit sanguin. Une pression artérielle trop haute entraîne des risques d'accidents vasculaires importants.

III.1.1.iii La matrice extracellulaire

La matrice extracellulaire (MEC) correspond à l'ensemble des macromolécules présentes dans l'espace extracellulaire. Ces macromolécules se divisent en quatre grands groupes : les collagènes, les fibres élastiques, les glycoprotéines de structure et les protéoglycanes.

(i) Les collagènes

Les collagènes sont les protéines les plus représentées dans la paroi vasculaire. Il en existe 19 différentes réparties en deux sous-familles : les collagènes de type fibrillaire (dont les types I, II et III) et les collagènes de type non fibrillaire (dont le type IV)¹⁴. Les collagènes de type fibrillaire s'assemblent pour constituer des fibrilles de collagènes. Ces dernières s'agrègent ensuite pour former les fibres de collagènes. Le collagène de type IV est le collagène le plus retrouvé au niveau de la lame basale où il forme un réseau en multicouche¹⁴. Sary a montré que 60% du volume de la plaque était composé de collagènes¹⁵. Il a également démontré, que bien que le collagène IV soit généralement présent au niveau de la lame basale, il pouvait entourer les CML de la média au cours de la progression de la plaque d'athérome. Enfin le collagène a un rôle de renfort et de stabilisation en opposition à l'action des métalloprotéinases libérées par les CML et par les macrophages présents au sein de la plaque¹⁶.

(ii) Les fibres élastiques

La présence de ces fibres permet d'assurer la fonction élastique des vaisseaux. Ces fibres sont composées d'élastine et de glycoprotéines à structure microfibrillaire comme la fibrilline 1 et 2 et les MAGP (microfibrillar-associated glycoprotein). Il a été montré qu'une activation du gène unique de l'élastine chez des souris engendrait une prolifération excessive des CML et une occlusion totale de l'artère¹⁷. De plus, ces fibres élastiques jouent un rôle important dans la fenestration présente au niveau des limitantes élastiques. En effet, elles permettent le passage de substances nutritives diffusant à travers les lames, de la lumière vasculaire vers la média. Elles permettent également d'établir des contacts cellule-cellule ou cellule-MEC entre l'intima et la média. Lors du vieillissement mais également de l'athérosclérose, il peut y avoir une rigidification de l'artère par augmentation de la quantité de Ca²⁺ liée aux fibres élastiques¹⁶.

(iii) Les glycoprotéines de structure

Les principales glycoprotéines sont la vitronectine, la fibronectine, la laminine, la thrombospondine et la tenascine. Leur séquence peptidique leur permet de s'auto-associer, de s'associer à d'autres protéines de la MEC mais également de se lier à des récepteurs comme les intégrines. La vitronectine, notamment, a été mise en évidence au sein des plaques d'athérome par immunohistochimie mais également par la sélection d'un anticorps dirigé contre cette protéine par l'équipe dans laquelle j'ai effectué mes trois années de thèse¹⁸.

(iv) Les protéoglycanes

A l'exception de l'acide hyaluronique, elles sont composées de longues chaînes de glycosaminoglycanes (GAG) liées de façon covalente à un noyau protéique. Leur structure à

la fois glucidique et protéique explique l'interaction importante avec les protéines matricielles et cellulaires. Des marquages immunohistochimiques ont montré que ces protéoglycanes se répartissent de manière différente suivant le stade de progression de la plaque. Par exemple, les hyaluronans et les versicans sont présents au stade précoce alors que les biglycans et les decorines le sont dans un stade plus avancé ¹⁹.

III.1.2 Acteurs cellulaires immuns

III.1.2.i *Monocytes et macrophages*

Les monocytes correspondent à des cellules mononucléées. Ils sont issus de promonocytes eux-mêmes provenant des cellules progénitrices myélomonocytaires. Les monocytes, une fois matures, vont être les précurseurs des macrophages mais aussi de certaines cellules dendritiques. Des marqueurs exprimés à la surface cellulaire ont permis de définir des sous populations fonctionnelles de monocytes (4 chez la souris et 2 chez l'humain). Les monocytes murins Ly6c^{hi} CXCR3^{low} GR1^{hi} CCR2^{hi} et humains CD14⁺⁺ CD16⁻ se caractérisent par leur aptitude à coloniser les tissus inflammatoires alors que les monocytes murins Ly6c^{low} CXCR3^{hi} GR1^{low} CCR2^{low} et humains CD14^{low} CD16⁺ sont les précurseurs des macrophages résidents ou des cellules dendritiques ^{20,21}. Il a été montré que les macrophages sont les cellules recrutées en premier lieu lors des premières étapes de formation de la plaque d'athérome. L'activation des cellules endothéliales par l'oxydation et la modification des lipoprotéines (LDLox et MDA-LDL (Malonaldehyde-modified low density lipoprotein), respectivement) engendre une sécrétion de chimiokines qui vont alors recruter les monocytes du sang ¹³. Faisant suite au chimiotactisme, les monocytes vont rouler sur les cellules endothéliales par interaction entre la PSGL1 (P-selectin glycoprotein ligand 1) et les sélectines exprimées par les cellules endothéliales. C'est « l'adhésion de roulement ». Une liaison irréversible s'établit ensuite entre VCAM-1 et ICAM-1 exprimées par les cellules endothéliales et VLA-4 (Very late antigen 4) et LFA-1 (Lymphocyte function-associated antigen), molécules exprimées par les monocytes, respectivement. C'est « l'adhésion ferme ». Il est à noter que les plaquettes jouent un rôle important dans l'interaction des monocytes aux cellules endothéliales par sécrétion de chimiokines et par activation de la voie NF-κB qui engendre l'expression de molécules d'adhésion ^{20,22}. S'ensuit la diapédèse, migration transendothéliale des monocytes. Les monocytes, au site inflammatoire, vont se différencier en macrophages, via le facteur M-CSF (Macrophage-Colony Stimulating Factor). Ils acquièrent alors la capacité à phagocyter les ApoB-LPs via les récepteurs SR-A (A scavenger receptor) et CD36. Les macrophages gorgés de lipides vont devenir des cellules dites « spumeuses » ^{23,24}. De plus, les macrophages chargés en molécules lipidiques sécrètent des protéoglycanes, comme les GAG capables de lier les LDL amplifiant le phénomène de rétention des lipides ^{25,26}. L'absorption de LDLox n'a pas de rétrocontrôle métabolique. Les macrophages recyclent les récepteurs aux LDLox amplifiant le processus d'accumulation de cholestérol sous forme de gouttelettes lipidiques (<http://jleoni.free.fr/contenu/pdf/Couv.pdf> - thèse de Jérôme LEONI). Les macrophages spumeux intra-plaque synthétisent des cytokines inflammatoires ²⁷ et des ROS issus du « burst » oxydatif augmentant la réponse inflammatoire ainsi que des MMPs détruisant la MEC et fragilisant la plaque. Les macrophages apoptotiques ne synthétisant plus de TGFβ, les CML, dépendantes de ce facteur, ne synthétisent plus de collagène (**Figure 4**) ²⁸.

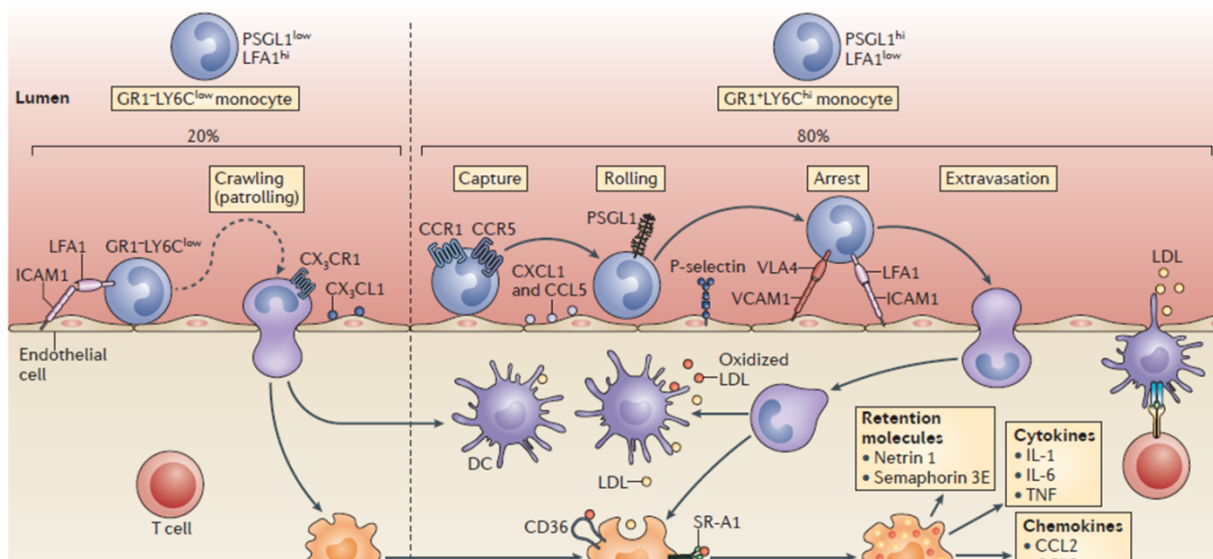


Figure 4 Mécanismes régulant le recrutement des monocytes et leur accumulation dans la plaque d'athérome. L'hyperlipidémie augmente le nombre de monocytes GR1 + LY6Chi, qui constituent 80% des monocytes recrutés chez la souris, le reste étant des monocytes de patrouille GR1^{LY6C}^{low}. Ces types de monocytes utilisent différentes paires de chimiokine/récepteur aux chimiokines pour infiltrer l'intima, ce qui est facilité par les molécules d'adhésion endothéliale, y compris les sélectines, ICAM-1 et VCAM-1. Les monocytes recrutés se différencient en macrophages ou cellules dendritiques (DC) dans l'intima, où ils interagissent avec des lipoprotéines athérogènes. Les macrophages phagocytent des LDL oxydées par la voie de la macropinocytose ou des voies induites par les récepteurs « scavengers » (y compris via le récepteur A1 (SR-A1) et le CD36), ce qui entraîne la formation de cellules spumeuses caractéristiques de la plaque. Ces cellules spumeuses sécrètent des cytokines pro-inflammatoires (l'IL-1, l'IL-6 et le TNF α), des chimiokines (CCL2, CCL5 et CX3C-CXCL1) ainsi que des facteurs de rétention des macrophages (netrin 1 et semaphorine 3E) amplifiant la réponse inflammatoire. CX3CL1, CX3C-chimokine ligand 1; CX3CR1, CX3C-chimokine récepteur 1; LFA1, antigène 1 associé à une fonction lymphocytaire; PSGL1, ligand 1 de glycoprotéine de P-sélectine; VLA4, antigène très tardif 4; IL, interleukine; TNF, Tumor necrosis factor (Extrait de l'article de Moore et al. paru en 2013 dans *Nature Review Immunology*).

III.1.2.ii Les lymphocytes T

(i) D'un point de vue historique

L'environnement en cytokines influence la différenciation des lymphocytes T en lymphocytes T auxiliaires. Un environnement riche en IL12 (Interleukin12) active le facteur T-bet entraînant la production d'IFN γ , qui va inhiber la production de l'IL-4 et orienter la différenciation des lymphocytes T vers Th1. Il en est de même pour IL-18. La population lymphocytaire Th2 se différencie en présence de l'IL4 qui augmente l'expression du facteur GATA-3 inhibant alors l'IFN γ (Interferon gamma). Les lymphocytes Th17 se différencient en présence de TGF β ainsi que d'IL6, d'IL21, d'IL23 et d'IL1 β et des facteurs tels que les facteurs de transcription STAT3, ROR γ T et ROR α ^{29,30}. L'IL10 et le TGF β induisent quant à eux une différenciation des lymphocytes T vers une population Treg. De plus il a été montré récemment que la population DC (Dendritic cells) CD103⁺ CD11c⁺ serait également responsable d'une orientation Treg ³¹. Chacune des sous-populations de lymphocytes T auxiliaires est rattachée à une fonction particulière avec un profil cytokinique unique. Les Th1 sont dites pro-inflammatoires avec sécrétion d'IFN γ . Les lymphocytes Th2 sont connus pour leur implication dans les allergies avec la production d'IL4, d'IL5 et d'IL6. Les Th17 avec la sécrétion d'IL17 et d'IL22 ont un rôle important dans les maladies auto-immunes. Les lymphocytes Treg sont impliqués dans la régulation de la réponse inflammatoire et le maintien de l'homéostasie en inhibant les autres populations de lymphocytes T effecteurs.

(ii) Les lymphocytes T dans l'athérosclérose

Dans l'athérosclérose, les cellules lymphocytaires T sont recrutées selon le même procédé que les macrophages. Le récepteur VCAM-1 des cellules endothéliales intervient dans la liaison des lymphocytes T CD4⁺ avant extravasation dans l'espace sous-endothélial grâce aux chimiokines telles que l'IP-10 (Interferon gamma induced protein - 10), l'ITAC (Interferon-inducible T-cell Alpha Chemoattractant) et Mig (Monokine induced by gamma interferon). Le rôle de chaque sous-population de lymphocytes TCD4⁺ citées plus haut a été étudié en particulier chez deux modèles de souris KO (Knockout) LDLR^{-/-} et ApoE^{-/-} (**Figure 5**). Ces deux modèles souris sont les modèles les plus utilisés dans l'athérosclérose. Leur déficience en ApoE et en récepteurs aux LDL (au niveau hépatique) fait que ces deux modèles développent des plaques d'athérome. Dans le cadre de l'athérosclérose, les Th1 prédominent par rapport aux Th2. Les macrophages activés dans la plaque vont sécréter de l'IL12 et de l'IL18 qui orientent la différenciation des cellules T vers Th1. L'IFN γ alors sécrété va conduire à une exacerbation de l'inflammation dans la plaque. Le modèle souris double KO ApoE^{-/-} IFN γ R^{-/-} a montré une réduction de 60% de la lésion athéromateuse³². De même des anticorps bloquant l'IL12 injectés chez des souris ApoE^{-/-} ont permis une réduction des lésions de 68%³³. Chez l'homme, un rôle athéro-protecteur des cellules Th2 a été mis en exergue³⁴. Un taux élevé de cellules lymphocytaires Th2 a été associé à une plus faible épaisseur de la paroi intimale intra carotidienne ainsi qu'une diminution du risque d'infarctus du myocarde chez la femme. Le rôle des lymphocytes Th17 dans le cadre de l'athérosclérose est mal étudié. Aucune donnée n'a été rapportée à ce jour. Les lymphocytes Treg présenteraient, quant à eux, un rôle athéro-protecteur par sécrétion d'IL10 et de TGF β , ainsi qu'un rôle important dans l'inhibition de la formation des cellules spumeuses^{35,36}. Cependant, il semblerait que les lymphocytes Treg soient peu représentés dans les plaques.

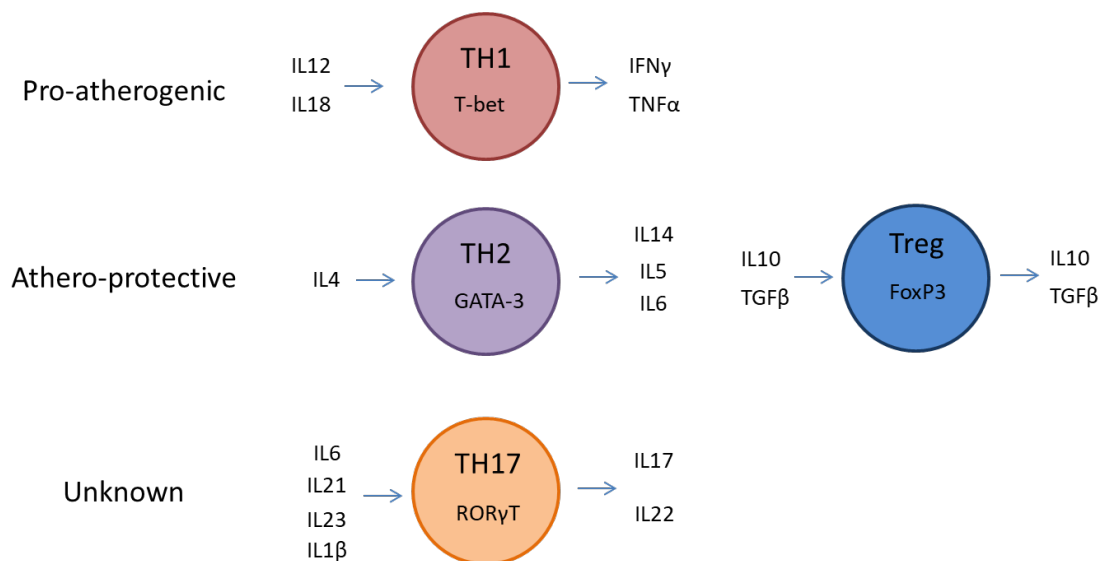


Figure 5 Les effets des sous-populations de cellules T sur l'athérosclérose. Pro-athérogène: Th1 (rouge). La présence d'IL12 et d'IL18 provoque une orientation vers Th1. IL12 entraîne l'activation du facteur de transcription, T-bet. L'activation de T-bet entraîne une production de cytokines pro-inflammatoires, y compris l'IFN γ et de TNF α . Athéro-protecteur: Treg (bleu). IL10 et TGF- β sont requis pour la différenciation des Treg. Les Treg expriment le facteur de transcription FoxP3 et peuvent produire de l'IL10 et / ou du TGF- β . TH2 (violet). Les cellules T CD4 + naïves s'engagent dans une lignée Th2 sous l'influence de l'IL4. Dans une boucle de réaction positive, IL-4 augmente l'expression du facteur de transcription GATA3, qui produit de l'IL4 et inhibe la production d'IFN- γ . Rôle inconnu : TH17 (orange). La différenciation Th17 nécessite une combinaison de plusieurs cytokines, y compris l'IL6, l'IL21, l'IL23 et l'IL1 β . Ces cellules expriment le facteur de transcription ROR γ T et produisent de l'IL17 et de l'IL22

En 2002, Caligiuri et *al* ont mis en évidence le rôle des lymphocytes B dans l'athérosclérose³⁷. Ces travaux ont montré que des souris ApoE^{-/-} ayant subi une splénectomie développaient des plaques plus importantes par comparaison à des souris ApoE^{-/-} qui ont conservé leur rate. Afin d'identifier le rôle des lymphocytes B par rapport aux lymphocytes T, ces deux populations cellulaires ont été séparées par séparation immunomagnétique. Le transfert de lymphocytes B chez des souris ayant subi une splénectomie a montré que ces cellules avaient un effet athéro-protecteur. Par ailleurs, une autre étude a montré qu'un anticorps anti-CD20 permettant la déplétion des lymphocytes B chez des souris LDLR^{-/-} ou ApoE^{-/-} engendrait une réduction de la plaque. Ces études ont donc suggéré l'existence, comme pour les lymphocytes T, de différentes sous-populations de lymphocytes B (**Figure 6**)³⁸. Ainsi, les deux sous-types B1 et B2 se distinguent d'après leur origine : les B1 proviennent de la cavité péritonéale et pleurale alors que les B2 proviennent de la moelle osseuse. Les B1 auraient un effet athéroprotecteur notamment par la production d'IgM qui inhibe la phagocytose des LDL par les macrophages et bloque donc la formation des cellules spumeuses^{39,40}. A contrario, les B2 seraient préférentiellement pro-athérogéniques⁴¹. En effet, une déplétion du récepteur BAFF (B-cell activating factor), exprimé spécifiquement chez les cellules B2, chez des souris BAFFR^{-/-} ApoE (Apolipoprotéine E) ^{-/-} entraîne une déplétion des lymphocytes B2 associée à une réduction de la plaque. Une nouvelle sous-population de lymphocytes B a également été découverte : les cellules B IRA (Innate Response Activator), sous-population de lymphocytes B1 répondant à une stimulation au LPS (lipopolysaccharide) et exprimant du GM-CSF (Granulocyte/Macrophage-Colony Stimulating Factor). Il semblerait que les lymphocytes B exprimant IRA soient pro-athérogéniques. Ces cellules B stimulent l'expansion des cellules dendritiques matures orientant une différenciation des cellules T vers un phénotype Th1 sécrétant de l'IFN γ . L'IFN γ produit entraîne un switch isotypique de l'IgG1 vers l'IgG2c en réponse aux LDL oxydées (**Figure 6**)⁴².

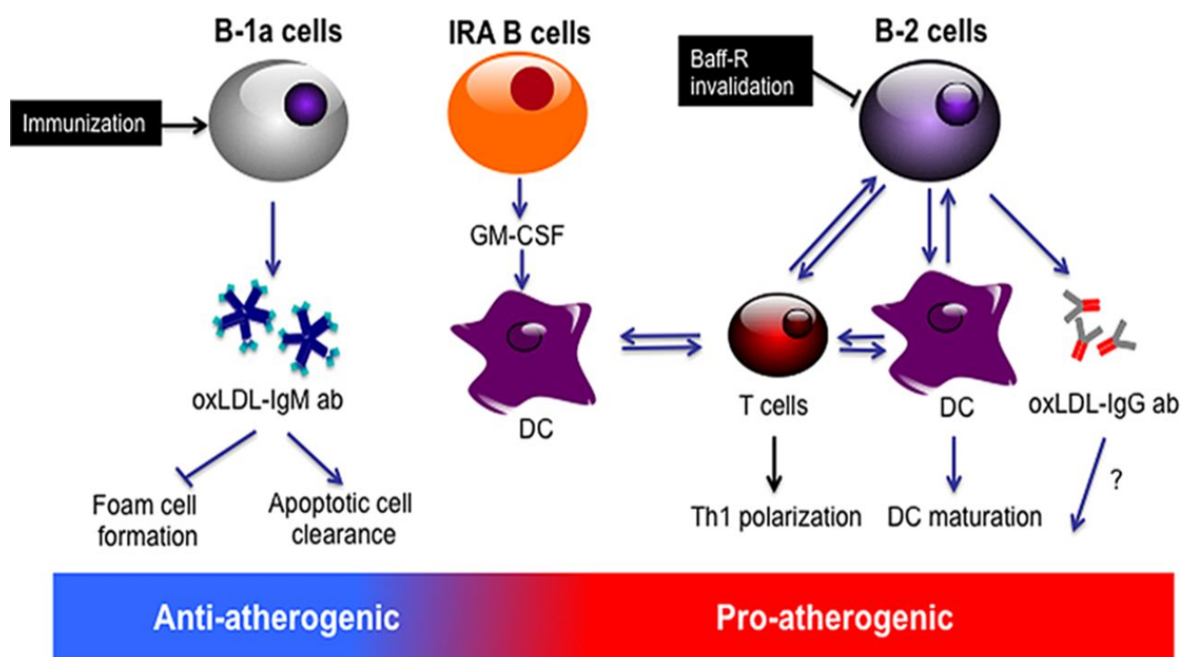


Figure 6 Réponses des lymphocytes B dans l'athérosclérose. Les cellules B-1a se sont avérées être athéroprotectrices principalement par la libération d'anticorps anti-LDLoxydés (oxLDL). Les cellules B-2 favorisent l'athérosclérose, stimulent l'activation des cellules T, des cellules dendritiques (DC) et la polarisation vers Th1. Les cellules B IRA sécrètent le facteur de stimulation des colonies de granulocytes-macrophages (GM-CSF), favorisent l'expansion DC et la polarisation des cellules Th1. On a montré que les stratégies thérapeutiques basées sur l'expansion B-1a (protocole d'immunisation) ou la déplétion de B-2 (invalidation de Baff-R) réduisaient l'athérosclérose dans des modèles expérimentaux. BAFF indique le facteur d'activation des cellules B; et BAFF-R, récepteur BAFF (Extrait de l'article de Ait-Oufella et al. paru en 2014 dans *Circ Res.*)

Les plaquettes ont un rôle prépondérant dans l'athérosclérose. Elles interviennent tout d'abord lors de la formation de la plaque d'athérome : les plaquettes, par la molécule PSGL-1 et la GPIb α vont se fixer à l'endothélium exprimant la P-sélectine (**Figure 7**). Ainsi, les plaquettes vont libérer des chimiokines (RANTES (Regulated upon Activation, Normal T-cell Expressed and Secreted ou CCL5) et le PF-4 (Platelet Factor-4 ou CXCL4)). Ces deux molécules vont attirer les monocytes ainsi que les LDLox. Le CD40L à la surface des plaquettes va, quant à lui, se fixer au CD40 exprimé par les cellules endothéliales aboutissant à la libération de MCP-1, de l'IL-8 et de protéases telles que les MMP⁴³. L'action combinée des LDL oxydées ainsi que des plaquettes activées va entraîner la production d'IL-1 β capable d'activer la voie NF- κ B (Nuclear factor – kappa B) des cellules endothéliales pour la synthèse de molécules de l'inflammation telles que ICAM-1, VCAM-1 qui participent au recrutement des leucocytes et donc à l'augmentation de la réponse inflammatoire^{44,45}. Les plaquettes vont ensuite intervenir, dans un deuxième temps, au niveau du thrombus⁴⁶. Lors de la rupture de la plaque, l'exposition de protéines de la MEC telles que le collagène ou vWF (Von Willebrand Factor) conduit au recrutement des plaquettes au site lésionel. Les plaquettes adhèrent à la MEC par la liaison de la glycoprotéine GPIb/IX/V à vWF et de GPVI au collagène ce qui engendre une modification conformationnelle et une activation de la GPIIb/IIIa (GP α Ib β 3), se liant au fibrinogène, et de la GP α Ib β 1 (ou GPIaIIa), se liant au collagène^{47,48}. La thrombine, générée rapidement, est considérée comme le facteur principal d'activation plaquettaire conduisant à l'activation des intégrines et à la dégranulation plaquettaire. L'ADP (Adenosine diphosphate) et la TXA2 (Thromboxane A2) sont libérées et entraînent un changement conformationnel de l'intégrine GP α Ib β 3 qui se fixe alors de manière ferme au fibrinogène. Il a été montré qu'un blocage de la TXA2 par des antagonistes tels que l'aspirine ou l'indométhacine engendre une diminution de la plaque d'athérome^{49,50,51}.

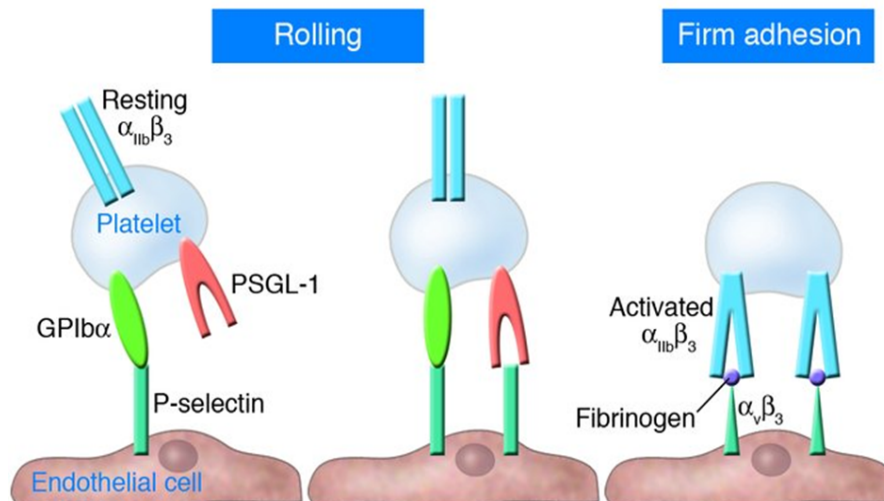


Figure 7 Adhésion plaquette-endothélium. La surface active de l'endothélium exprime la P-sélectine. Les récepteurs de surface plaquettaire GPIb α et PSGL-1 interagissent avec la P-sélectine endothéliale et provoquent le roulement des plaquettes. L'adhérence ferme ultérieure est médiée par les intégrines β_3 (Extrait de l'article de Gawaz et al. paru en 2005 dans *JCI*).

III.2 Facteurs moléculaires

III.2.1 Le cholestérol

Les études ont montré que des taux élevés de LDL cholestérol ainsi que d'ApoB, composant majeur des LDL seraient directement associés aux risque d'évènements cardiovasculaires. En effet, comme décrit dans la section « athérogénèse », l'infiltration et la rétention d'ApoB dans

la paroi artérielle est l'évènement initiateur qui engendre une réponse inflammatoire puis le développement de la plaque d'athérome. L'oxydation d'ApoB va provoquer une activation endothéliale suivie d'un recrutement cellulaire et d'une réponse immune exacerbée. Contrecarrant cet effet, les HDL (constitués de l'ApoA1 et l'ApoE) augmentent l'efflux du cholestérol présent dans les macrophages. Ainsi, la réponse inflammatoire et le stress oxydatif liés aux macrophages seront atténués. En effet, les HDL auraient un rôle protecteur en contrôlant l'efflux de cholestérol par les macrophages spumeux vers le foie pour excrétion, processus appelé transport inverse de cholestérol. Ceci est permis grâce à une régulation positive des deux récepteurs : ABCA1 (ATP-cassette subfamily A1) et ABCG1 (ATP-cassette subfamily G1). La transcription des gènes codant pour ces deux protéines est contrôlée par les récepteurs LXR (Liver X receptor), récepteurs nucléaires activés par les oxystérols, c'est-à-dire les formes oxydées du cholestérol. Brièvement, le foie charge les lipides sur les ApoB et sécrètent des lipoprotéines de très faible densité (VLDLs). Ces VLDLs vont former des LDLs après lipolyse par la lipoprotéine lipase. Les LDLs ainsi formées vont interagir avec les récepteurs aux LDLs (LDLR) du foie et être absorbées. Les lipoprotéines de forte densité (HDLs) sont quant à elles produites par l'intestin et le foie par la sécrétion d'ApoA1 (Apolipoprotéine A1) libres de lipides. L'ApoA1 va alors recruter le cholestérol grâce au transporteur ABCA1 formant des HDLs naissants. Dans les organes périphériques, les HDLs naissants vont permettre l'efflux de cholestérol par l'intermédiaire d'ABCA1. Les HDLs dit matures après estérification grâce à l'enzyme lécithine cholestérol acétyltransférase (LCAT) vont également permettre cet efflux mais par l'intermédiaire d'un autre transporteur dit ABCG1. Le cholestérol dans les HDLs est redirigé vers le foie par deux voies : l'absorption directe des HDL matures par les récepteurs SR-B1 ou suite à leur transformation vers les LDLs et VLDLs via l'action de la protéine de transfert d'ester de cholesteryl (CETP) (**Figure 8**).

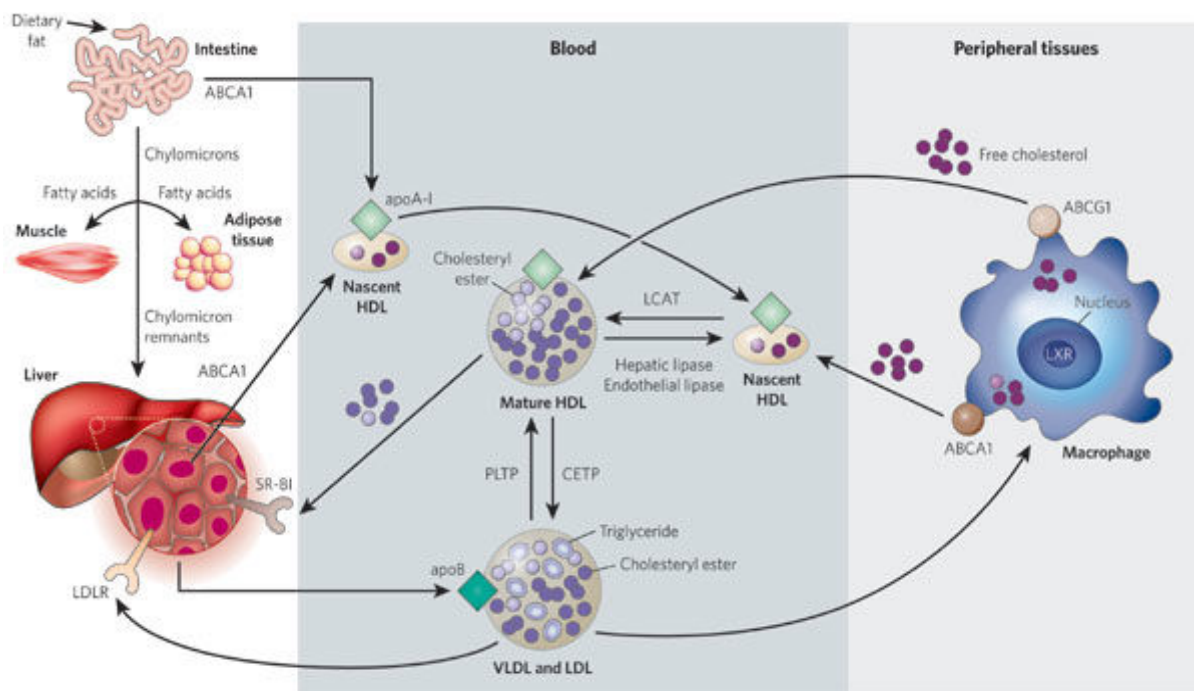


Figure 8 Le métabolisme des lipoprotéines joue un rôle clé dans l'athérogénèse. Il implique le transport de lipides, en particulier le cholestérol et les triglycérides, dans le sang. L'intestin absorbe les graisses alimentaires et les conditionne en chylomicrons (grandes lipoprotéines riches en triglycérides), qui sont transportés vers les tissus périphériques par le sang. Dans les tissus musculaires et adipeux, l'enzyme lipoprotéine lipase décompose les chylomicrons, et les acides gras pénètrent dans ces tissus. Les restes de chylomicrons sont ensuite absorbés par le foie. Le foie charge les lipides sur l'apoB et sécrète des lipoprotéines de très basse densité (VLDL), qui subissent une lipolyse par la lipoprotéine lipase pour former des lipoprotéines de basse densité (LDL). Les LDL sont ensuite absorbées par le foie par liaison au récepteur des LDL (LDLR), ainsi que par d'autres voies. En revanche, les lipoprotéines de haute densité (HDL) sont générées par l'intestin et le foie par la sécrétion d'apoA-I sans lipides. L'ApoA-I recrute ensuite le cholestérol à partir de ces organes par les actions du transporteur ABCA1, formant des HDL naissants, ce qui protège l'apoA-I d'une dégradation rapide dans les reins. Dans les tissus périphériques, les HDL naissants favorisent l'efflux de cholestérol des tissus, y compris des macrophages, par l'action de ABCA1. Les HDL matures favorisent également cet efflux, mais à travers l'action d'ABCG1. (Dans les macrophages, le récepteur nucléaire LXR régule positivement la production d'ABCA1 et d'ABCG1.) Le cholestérol libre (non estérifié) dans les HDL naissants est estérifié en cholestéryl ester par l'enzyme lécithine cholestérol acyltransférase (LCAT). Le cholestérol dans les HDL est renvoyé au foie à la fois directement par absorption par le récepteur SR-BI et indirectement par transfert aux LDL et aux VLDL par la protéine de transfert d'ester de cholestéryle (CETP). La teneur en lipides des HDL est altérée par les enzymes lipase hépatique et endothéliale et par les protéines de transfert CETP et la protéine de transfert des phospholipides (PLTP), affectant le catabolisme des HDL (Extrait de l'article de Daniel J. Rader & Alan Daugherty, paru en 2008 dans *Nature*)

Ainsi plusieurs mécanismes peuvent être ciblés pour réduire le taux de cholestérol notamment stimuler le métabolisme des HDLs par injection d'ApoA1, en utilisant des agonistes des LXR⁵² ainsi que des polyphénols comme le resvératrol⁵³ qui agissent sur la voie PPAR (peroxisome proliferator-activated receptor) – LXR – ABCA1. Il a été également montré que la protéine PCSK9 (proprotéine convertase subtilisine/kexine de type 9) impliquée dans la diminution de l'expression de récepteurs aux LDL au niveau du foie pouvait être ciblée par des anticorps monoclonaux anti PCSK9 produits par Sanofi (alirocumab) et Amgen (evolocumab). Ainsi, ces anticorps bloquant la PCSK9 permettent indirectement une réduction du taux de LDL-cholestérol dans le sang. ApoB fait également partie des cibles par utilisation d'anticorps anti-ApoB⁵⁴ ou d'inhibiteurs de ApoB (utilisation d'ARN interférents)⁵⁵. Actuellement, lorsqu'un patient présente un taux anormalement haut de cholestérol, les statines sont utilisés afin de réduire le taux de VLDLs, de LDLs et de triglycérides et d'augmenter le taux de HDL.

III.2.2 Les espèces réactives oxygénées

Les trois espèces réactives oxygénées (ROS) les plus connues sont l'anion superoxyde, le peroxyde d'hydrogène et le radical hydroxyle. Les ROS sont générées au sein de la paroi du vaisseau par plusieurs types cellulaires comme les CML ou les cellules endothéliales mais également par les macrophages, les lymphocytes et les plaquettes^{56,57}. Cette production se fait par plusieurs mécanismes, notamment la voie de la NADPH (Nicotinamide adénine dinucléotide phosphate) oxydase^{57,58}. La formation des ROS peut être stimulée par un stress mécanique, les facteurs environnants, les cytokines, l'angiotensine II, les AGE (Advanced glycation end-products) et les LDL natifs⁵⁹. Les ROS contrôlent plusieurs fonctions notamment la croissance cellulaire ainsi que la prolifération et le contrôle de la vasoconstriction et la vasodilatation des vaisseaux⁵⁷. Ils sont impliqués dans l'oxydation des LDL⁵⁹ et dans la diminution de NO qui engendre une dysfonction des cellules endothéliales⁶⁰. Un autre effet athérogène du peroxyde d'hydrogène est l'activation du facteur de transcription NF- κ B et de AP-1 (Activator protein 1) induisant l'expression de molécules d'adhésion telles que VCAM, ICAM-1, la E-selectine et d'autres cytokines (**Figure 9**). Les ROS augmentent l'expression de divers facteurs de croissance et activent divers signaux de stress tels que JNK (c-Jun N-terminal kinases) ou Pim-1 (Proto-oncogene serine/threonine-protein kinase), ce qui conduit à la prolifération des CML⁶⁰.

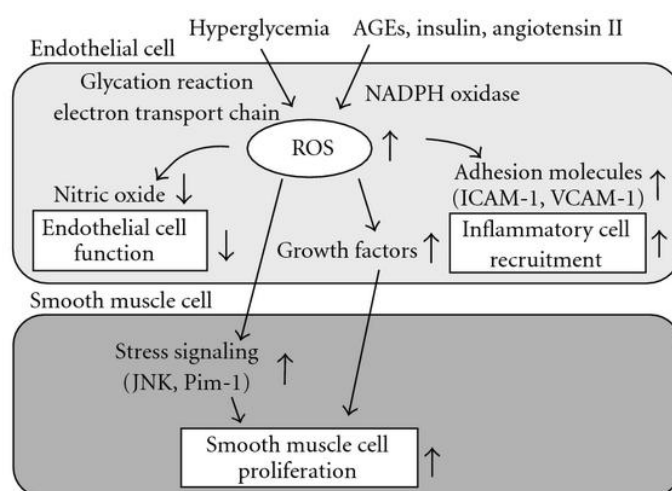


Figure 9 Rôle des ROS dans le développement de l'athérosclérose. Les ROS sont produits par diverses voies dans des conditions diabétiques et impliqués dans le développement de l'athérosclérose sous divers aspects. L'hyperglycémie induit la production de ROS par activation de la réaction de glycation et de la chaîne de transport des électrons dans les mitochondries. En outre, les AGE (Advanced glycation end-product), l'insuline et l'angiotensine II induisent la production de ROS par l'activation de la NADPH oxydase. Les ROS sont impliqués dans le développement de l'athérosclérose à différents niveaux. Tout d'abord, la production de ROS diminue les niveaux d'oxyde nitrique (NO), ce qui conduit à un dysfonctionnement des cellules endothéliales. Deuxièmement, les ROS augmentent l'expression de diverses molécules d'adhésion telles que ICAM-1 et VCAM-1, ce qui conduit à un recrutement cellulaire inflammatoire. Enfin, les ROS augmentent l'expression de divers facteurs de croissance et activent divers signaux de stress tels que JNK et Pim-1, ce qui conduit à la prolifération des cellules musculaires lisses (Extrait de l'article de Kaneto et al. paru en 2010 dans *Mediators Inflamm.*)

III.2.3 Les cytokines

Le terme « cytokine » a été introduit en 1974 par Stanley Cohen⁶¹. Il s'agit de substances solubles impliquées dans la signalisation cellulaire et synthétisées par les cellules du système immunitaire ou par d'autres cellules ou tissus, de manière à réguler à distance la fonction et l'activité cellulaire. Chaque type cellulaire est en capacité de répondre mais également de produire des cytokines. Les cytokines sont classées en plusieurs catégories : les interleukines (IL), les facteurs de nécrose tumorale (TNF), les chimiokines, les interférons (IFN), les facteurs stimulant l'hématopoïèse (CSF) et les facteurs de croissance de transformation

(TGF)⁶². Au cours de l'athérogénèse, les cytokines ont un rôle prépondérant et contrôlent l'évolution de la plaque. Les cytokines ont été étudiées dans le cadre de l'athérosclérose dans deux modèles animaux, les souris ApoE -/- et les souris LDLR -/-^{63,64}. Au stade précoce de l'athérogénèse, l'endothélium présente une lésion qui le rend perméable aux molécules telles que les cytokines qui vont réguler l'activité cellulaire. Au niveau de la barrière endothéliale, l'IFN γ et le TNF α vont réguler l'ouverture entre cellules par réorganisation de l'actine et de la tubuline⁶⁵. Les cellules endothéliales vont alors exprimer des molécules d'adhésion mais également sécréter des chimiokines, impliquées dans le recrutement des différents types cellulaires. Les couples chimiokines/récepteur aux chimiokines les plus impliqués dans le recrutement cellulaire et notamment des monocytes sont CCL2/CCR2, CX₃CL1/CX₃CR1 et CCL5/CCR5⁶⁶. Chaque chimiokine possède des effets pro- ou anti- athérogènes. Le tableau ci-dessous récapitule l'action de chacune (**Tableau 1**)⁶⁷.

Tableau 1. Rôle pro- ou anti- athérogène des chimiokines

Chimiokines pro-athérogènes	Chimiokines anti-athérogènes
CCL2 (MCP-1), CCL3 (MIP-1 α), CCL5 (RANTES), CXCL10 (IP-10), CX3CL1, CCL17, CXCL4 (PF-4), MIF, Mig (CXCL9), ITAC (CXCL11)	CXCL5, CXCL1, CXCL16, CCL19/CCL21, CXCL12,

Complémentant le rôle attracteur de ces molécules, les chimiokines peuvent avoir des fonctions autres. Par exemple, CX3CL1 est connu pour son rôle de molécule d'adhésion⁶⁸. CCL2 ou MCP-1 entraîne une baisse de transport inverse du cholestérol par répression de l'expression des protéines impliquées dans l'efflux de cholestérol telles que ABCA1, ABCG1 et SR-B1⁶⁹. En médecine, les statines, bien que connues pour leur rôle sur le taux de cholestérol, peuvent être également utilisées afin de modifier les couples chimiokines/récepteur aux chimiokines de manière à atténuer l'expression de ces derniers⁷⁰. D'autres approches sur les modèles animaux ont été tentées comme la sialylation qui perturbent l'interaction chimiokines/récepteur aux chimiokines⁷¹, l'utilisation d'antagonistes/d'inhibiteurs spécifiques des récepteurs⁷², le « gene silencing » par utilisation de nanoparticules contenant des ARN interférents⁷³, l'inhibition des chémokines par transfert de gènes grâce aux lentivirus^{74,75} ou encore les anticorps bloquants^{76,77}. Suite à l'attraction des leucocytes, une réaction immune se met en place impliquant notamment les interleukines qui vont influencer l'évolution de la lésion athéromateuse. Le tableau suivant donne le rôle de chacune (**Tableau 2**).

Tableau 2. Rôle pro- ou anti- inflammatoire des interleukines

Interleukines pro-inflammatoires	Interleukines anti-inflammatoires
IL1 α , IL1 β , IL2, IL6, IL12, IL15, IL18, IL20, IFN γ , TNF α	IL3, IL4, IL5, IL10, IL13, IL17, IL19, IL25, IL27, IL33, IL22, IL23, TGF β

Les interleukines ainsi que d'autres cytokines ont le pouvoir d'influencer la différenciation des macrophages qui va agir sur la formation des cellules spumeuses. Les macrophages suivent une dichotomie parallèle à celle des TH1/TH2 avec les phénotypes pro-inflammatoires (M1 sécrétant de l'IL6, l'IL12 et du TNF α) et immunorégulateurs (M2 produisant du TGF β et de l'IL-10)⁷⁸⁻⁸¹. L'environnement en cytokine va induire une orientation vers l'un ou vers l'autre des deux phénotypes. L'IL1 β et l'IFN γ , cytokines produits par les lymphocytes T Th1, vont conduire à une différenciation M1 alors que l'IL4 et IL13, cytokines produites par les lymphocytes Th2, vont induire une différenciation vers le phénotype M2 qui possède plutôt un rôle anti-athérogène⁷⁹⁻⁸¹.

III.2.4 Les métalloprotéinases de la matrice (MMP)

Les MMPs sont des protéases avec une activité protéolytique. Les métalloprotéinases peuvent avoir un rôle régulateur comme délétère. En effet, elles peuvent dégrader la MEC et fragiliser la plaque ou bien être impliquées dans le recrutement des cellules musculaires lisses et participer à la stabilisation de la plaque. Parmi les MMPs surexprimées dans la lésion, il a été montré une présence de MMP3, MMP1, MMP9 et MMP8. Il a également été démontré l'expression de MMP8, 11, 14 et 13 dans les zones de rupture de plaque. Les métalloprotéines 1, 2, 8, 13 et 14 ont la capacité de cliver le collagène. Les MMPs 9 et 2 fragmentent l'élastine et le collagène de type IV suggérant un rôle dans la rupture de la plaque⁸². Cependant certaines métalloprotéinases ont des rôles contradictoires comme MMP9. Une étude récente sur des souris double KO MMP9/ApoE a montré que MMP9 avait un rôle athéro-protecteur et stabilisateur au niveau de la plaque⁸³. En effet, les souris déplétées de MMP9 présentaient des plaques plus importantes au niveau de l'artère brachio-céphalique. Parmi les protéinases se trouvent également les cathepsines qui sont impliquées dans la formation de la plaque d'athérome et la déstabilisation. Il a été montré dans une étude que la cathepsine D était impliquée dans la modification des LDL augmentant leur reconnaissance par les macrophages et conduisant à la formation de cellules spumeuses⁸⁴. Un modèle souris déficient en cathepsines a montré le rôle de ces molécules dans l'infiltration des leucocytes, une dégradation de la lamina élastique, une néovascularisation ainsi qu'une rupture de la chape fibreuse⁸⁵.

III.3 Conclusions

Comme vu dans la section précédente III.2, une multitude d'acteurs cellulaires et de facteurs moléculaires, décrits dans la littérature comme étant présents et surexprimés dans l'athérome donne accès à de nombreux biomarqueurs pour le ciblage et le diagnostic de la plaque d'athérome. Plus récemment, comme cela sera décrit dans la partie « Résultats », une étude protéomique sur protéines athéromateuses issues de lapins athéromateux et de protéines saines issues de lapins sains a permis de mettre en exergue un certain nombre de marqueurs supplémentaires pouvant également être désignés comme de nouvelles cibles pour le développement d'objets ciblant fonctionnalisés par des anticorps humains pour une meilleure appréhension de l'athérogénèse et le diagnostic à différents stades de la plaque d'athérome.

Bibliographie

1. Insull, W. The Pathology of Atherosclerosis: Plaque Development and Plaque Responses to Medical Treatment. *Am. J. Med.* **122**, S3–S14 (2009).
2. Nording, H. M., Seizer, P. & Langer, H. F. Platelets in Inflammation and Atherogenesis. *Front. Immunol.* **6**, (2015).
3. Page, C. & Pitchford, S. Neutrophil and platelet complexes and their relevance to neutrophil recruitment and activation. *Int. Immunopharmacol.* **17**, 1176–1184 (2013).
4. Tabas, I. *et al.* Lipoprotein lipase and sphingomyelinase synergistically enhance the association of atherogenic lipoproteins with smooth muscle cells and extracellular matrix. A possible mechanism for low density lipoprotein and lipoprotein(a) retention and macrophage foam cell formation. *J. Biol. Chem.* **268**, 20419–20432 (1993).
5. Virmani, R., Kolodgie, F. D., Burke, A. P., Farb, A. & Schwartz, S. M. Lessons From Sudden Coronary Death. *Arterioscler. Thromb. Vasc. Biol.* **20**, 1262–1275 (2000).
6. Linton, M. F. *et al.* The Role of Lipids and Lipoproteins in Atherosclerosis. in *Endotext* (eds. De Groot, L. J. *et al.*) (MDText.com, Inc., 2000).
7. Kolodgie, F. D. *et al.* Intraplaque Hemorrhage and Progression of Coronary Atheroma. *N. Engl. J. Med.* **349**, 2316–2325 (2003).
8. Virmani, R. *et al.* Atherosclerotic Plaque Progression and Vulnerability to Rupture. *Arterioscler. Thromb. Vasc. Biol.* **25**, 2054–2061 (2005).
9. Libby, P. Inflammation in atherosclerosis. *Nature* **420**, 868–874 (2002).
10. Stary, H. C. *et al.* A Definition of Advanced Types of Atherosclerotic Lesions and a Histological Classification of Atherosclerosis: A Report From the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association. *Circulation* **92**, 1355–1374 (1995).
11. Malek, A. M. & Izumo, S. Mechanism of endothelial cell shape change and cytoskeletal remodeling in response to fluid shear stress. *J. Cell Sci.* **109** (Pt 4), 713–726 (1996).

12. cellule musculaire lisse vasculaire. Available at: <http://www.snv.jussieu.fr/bmedia/cmlv/cmlv-c2.htm>. (Accessed: 13th July 2017)
13. Gryglewski, R. J., Dembínska-Kieć, A. & Korbut, R. A possible role of thromboxane A₂ (TXA₂) and prostacyclin (PGI₂) in circulation. *Acta Biol. Med. Ger.* **37**, 715–723 (1978).
14. Barnes, M. J. & Farndale, R. W. Collagens and atherosclerosis. *Exp. Gerontol.* **34**, 513–525 (1999).
15. Stary, H. C. Natural history and histological classification of atherosclerotic lesions: an update. *Arterioscler. Thromb. Vasc. Biol.* **20**, 1177–1178 (2000).
16. Katsuda, S. & Kaji, T. Atherosclerosis and extracellular matrix. *J. Atheroscler. Thromb.* **10**, 267–274 (2003).
17. Urbán, Z. *et al.* Supravalvular aortic stenosis: genetic and molecular dissection of a complex mutation in the elastin gene. *Hum. Genet.* **109**, 512–520 (2001).
18. Robert, R. *et al.* Identification of human scFvs targeting atherosclerotic lesions: selection by single round in vivo phage display. *J. Biol. Chem.* **281**, 40135–40143 (2006).
19. Kolodgie, F. D. *et al.* Differential accumulation of proteoglycans and hyaluronan in culprit lesions: insights into plaque erosion. *Arterioscler. Thromb. Vasc. Biol.* **22**, 1642–1648 (2002).
20. Mestas, J. & Ley, K. Monocyte-Endothelial Cell Interactions in the Development of Atherosclerosis. *Trends Cardiovasc. Med.* **18**, 228–232 (2008).
21. Dufresne, S. S., Frenette, J. & Dumont, N. A. Inflammation et régénération musculaire - Une arme à double tranchant. *médecine/sciences* **32**, 591–597 (2016).
22. Koenen, R. R. *et al.* Disrupting functional interactions between platelet chemokines inhibits atherosclerosis in hyperlipidemic mice. *Nat. Med.* **15**, 97–103 (2009).
23. Johnson, J. L. & Newby, A. C. Macrophage heterogeneity in atherosclerotic plaques. *Curr. Opin. Lipidol.* **20**, 370–378 (2009).

24. Kunjathoor, V. V. *et al.* Scavenger receptors class A-I/II and CD36 are the principal receptors responsible for the uptake of modified low density lipoprotein leading to lipid loading in macrophages. *J. Biol. Chem.* **277**, 49982–49988 (2002).
25. Lindholm, M. W., Nilsson, J. & Moses, J. Low density lipoprotein stimulation of human macrophage proteoglycan secretion. *Biochem. Biophys. Res. Commun.* **328**, 455–460 (2005).
26. Maor, I., Hayek, T., Hirsh, M., Iancu, T. C. & Aviram, M. Macrophage-released proteoglycans enhance LDL aggregation: studies in aorta from apolipoprotein E-deficient mice. *Atherosclerosis* **150**, 91–101 (2000).
27. Moore, K. J., Sheedy, F. J. & Fisher, E. A. Macrophages in atherosclerosis: a dynamic balance. *Nat. Rev. Immunol.* **13**, 709–721 (2013).
28. Fadok, V. A. *et al.* Macrophages that have ingested apoptotic cells in vitro inhibit proinflammatory cytokine production through autocrine/paracrine mechanisms involving TGF-beta, PGE2, and PAF. *J. Clin. Invest.* **101**, 890–898 (1998).
29. Dong, C. Diversification of T-helper-cell lineages: finding the family root of IL-17-producing cells. *Nat. Rev. Immunol.* **6**, 329–334 (2006).
30. Korn, T., Bettelli, E., Oukka, M. & Kuchroo, V. K. IL-17 and Th17 Cells. *Annu. Rev. Immunol.* **27**, 485–517 (2009).
31. Subramanian, M., Thorp, E., Hansson, G. K. & Tabas, I. Treg-mediated suppression of atherosclerosis requires MYD88 signaling in DCs. *J. Clin. Invest.* **123**, 179–188 (2013).
32. Whitman, S. C., Ravisankar, P., Elam, H. & Daugherty, A. Exogenous Interferon- γ Enhances Atherosclerosis in Apolipoprotein E $^{-/-}$ Mice. *Am. J. Pathol.* **157**, 1819–1824 (2000).
33. Hauer, A. D. *et al.* Blockade of Interleukin-12 Function by Protein Vaccination Attenuates Atherosclerosis. *Circulation* **112**, 1054–1062 (2005).
34. Engelbertsen, D. *et al.* T-Helper 2 Immunity Is Associated With Reduced Risk of Myocardial Infarction and Stroke. *Arterioscler. Thromb. Vasc. Biol.* **33**, 637–644 (2013).

35. Mallat, Z. *et al.* Induction of a Regulatory T Cell Type 1 Response Reduces the Development of Atherosclerosis in Apolipoprotein E-Knockout Mice. *Circulation* **108**, 1232–1237 (2003).
36. Lin, J. *et al.* The role of CD4+CD25+ regulatory T cells in macrophage-derived foam-cell formation. *J. Lipid Res.* **51**, 1208–1217 (2010).
37. Caligiuri, G., Nicoletti, A., Poirier, B. & Hansson, G. K. Protective immunity against atherosclerosis carried by B cells of hypercholesterolemic mice. *J. Clin. Invest.* **109**, 745–753 (2002).
38. Ait-Oufella, H. *et al.* B cell depletion reduces the development of atherosclerosis in mice. *J. Exp. Med.* **207**, 1579–1587 (2010).
39. Ait-Oufella, H., Sage, A. P., Mallat, Z. & Tedgui, A. Adaptive (T and B Cells) Immunity and Control by Dendritic Cells in Atherosclerosis. *Circ. Res.* **114**, 1640–1660 (2014).
40. Binder, C. J. *et al.* Thematic review series: The Immune System and Atherogenesis. The role of natural antibodies in atherogenesis. *J. Lipid Res.* **46**, 1353–1363 (2005).
41. Tsiantoulas, D., Diehl, C. J., Witztum, J. L. & Binder, C. J. B Cells and Humoral Immunity in Atherosclerosis. *Circ. Res.* **114**, 1743–1756 (2014).
42. Hilgendorf, I. *et al.* Innate response activator B cells aggravate atherosclerosis by stimulating T helper-1 adaptive immunity. *Circulation* **129**, 1677–1687 (2014).
43. Huo, Y. *et al.* Circulating activated platelets exacerbate atherosclerosis in mice deficient in apolipoprotein E. *Nat. Med.* **9**, 61–67 (2003).
44. Henn, V. *et al.* CD40 ligand on activated platelets triggers an inflammatory reaction of endothelial cells. *Nature* **391**, 591–594 (1998).
45. Hawrylowicz, C. M., Santoro, S. A., Platt, F. M. & Unanue, E. R. Activated platelets express IL-1 activity. *J. Immunol.* **143**, 4015–4018 (1989).
46. Fitzgerald, D. J., Roy, L., Catella, F. & FitzGerald, G. A. Platelet Activation in Unstable Coronary Disease. *N. Engl. J. Med.* **315**, 983–989 (1986).

47. Arya, M. *et al.* Glycoprotein Ib–IX-mediated activation of integrin α IIb β 3: effects of receptor clustering and von Willebrand factor adhesion. *J. Thromb. Haemost.* **1**, 1150–1157 (2003).
48. Nieswandt, B. & Watson, S. P. Platelet-collagen interaction: is GPVI the central receptor? *Blood* **102**, 449–461 (2003).
49. Cayatte, A. J. *et al.* The Thromboxane Receptor Antagonist S18886 but Not Aspirin Inhibits Atherogenesis in Apo E–Deficient Mice. *Arterioscler. Thromb. Vasc. Biol.* **20**, 1724–1728 (2000).
50. Cyrus, T. *et al.* Effect of Low-Dose Aspirin on Vascular Inflammation, Plaque Stability, and Atherogenesis in Low-Density Lipoprotein Receptor–Deficient Mice. *Circulation* **106**, 1282–1287 (2002).
51. Praticò, D., Tillmann, C., Zhang, Z.-B., Li, H. & FitzGerald, G. A. Acceleration of atherogenesis by COX-1-dependent prostanoid formation in low density lipoprotein receptor knockout mice. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3358–3363 (2001).
52. Katz, A. *et al.* Safety, pharmacokinetics, and pharmacodynamics of single doses of LXR-623, a novel liver X-receptor agonist, in healthy participants. *J. Clin. Pharmacol.* **49**, 643–649 (2009).
53. Berrougui, H., Grenier, G., Loued, S., Drouin, G. & Khalil, A. A new insight into resveratrol as an atheroprotective compound: inhibition of lipid peroxidation and enhancement of cholesterol efflux. *Atherosclerosis* **207**, 420–427 (2009).
54. Schiopu, A. *et al.* Recombinant human antibodies against aldehyde-modified apolipoprotein B-100 peptide sequences inhibit atherosclerosis. *Circulation* **110**, 2047–2052 (2004).
55. Kastelein, J. J. P. *et al.* Potent reduction of apolipoprotein B and low-density lipoprotein cholesterol by short-term administration of an antisense inhibitor of apolipoprotein B. *Circulation* **114**, 1729–1735 (2006).

56. Madamanchi, N. R., Vendrov, A. & Runge, M. S. Oxidative stress and vascular disease. *Arterioscler. Thromb. Vasc. Biol.* **25**, 29–38 (2005).
57. Li, H., Horke, S. & Förstermann, U. Vascular oxidative stress, nitric oxide and atherosclerosis. *Atherosclerosis* **237**, 208–219 (2014).
58. Harrison, D., Griendling, K. K., Landmesser, U., Hornig, B. & Drexler, H. Role of oxidative stress in atherosclerosis. *Am. J. Cardiol.* **91**, 7A–11A (2003).
59. Mügge, A. The role of reactive oxygen species in atherosclerosis. *Z. Kardiol.* **87**, 851–864 (1998).
60. Kaneto, H., Katakami, N., Matsuhisa, M. & Matsuoka, T. Role of Reactive Oxygen Species in the Progression of Type 2 Diabetes and Atherosclerosis. *Mediators Inflamm.* **2010**, (2010).
61. Cohen, S. Cytokine: more than a new word, a new concept proposed by Stanley Cohen thirty years ago. *Cytokine* **28**, 242–247 (2004).
62. Tedgui, A. & Mallat, Z. Cytokines in atherosclerosis: pathogenic and regulatory pathways. *Physiol. Rev.* **86**, 515–581 (2006).
63. Getz, G. S. & Reardon, C. A. Animal models of atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **32**, 1104–1115 (2012).
64. Zadelaar, S. *et al.* Mouse models for atherosclerosis and pharmaceutical modifiers. *Arterioscler. Thromb. Vasc. Biol.* **27**, 1706–1721 (2007).
65. Pober, J. S. & Sessa, W. C. Evolving functions of endothelial cells in inflammation. *Nat. Rev. Immunol.* **7**, 803–815 (2007).
66. Tacke, F. *et al.* Monocyte subsets differentially employ CCR2, CCR5, and CX3CR1 to accumulate within atherosclerotic plaques. *J. Clin. Invest.* **117**, 185–194 (2007).
67. Ramji, D. P. & Davies, T. S. Cytokines in atherosclerosis: Key players in all stages of disease and promising therapeutic targets. *Cytokine Growth Factor Rev.* **26**, 673–685 (2015).

68. Teupser, D. *et al.* Major reduction of atherosclerosis in fractalkine (CX3CL1)-deficient mice is at the brachiocephalic artery, not the aortic root. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17795–17800 (2004).
69. Huang, C.-X., Zhang, Y.-L., Wang, J.-F., Jiang, J.-Y. & Bao, J.-L. MCP-1 impacts RCT by repressing ABCA1, ABCG1, and SR-BI through PI3K/Akt posttranslational regulation in HepG2 cells. *J. Lipid Res.* **54**, 1231–1240 (2013).
70. Nie, P. *et al.* Atorvastatin improves plaque stability in ApoE-knockout mice by regulating chemokines and chemokine receptors. *PloS One* **9**, e97009 (2014).
71. Döring, Y. *et al.* Deficiency of the Sialyltransferase St3Gal4 Reduces Ccl5-Mediated Myeloid Cell Recruitment and Arrest. *Circ. Res.* **114**, 976–981 (2014).
72. Olzinski, A. R. *et al.* Pharmacological inhibition of C-C chemokine receptor 2 decreases macrophage infiltration in the aortic root of the human C-C chemokine receptor 2/apolipoprotein E^{-/-} mouse: magnetic resonance imaging assessment. *Arterioscler. Thromb. Vasc. Biol.* **30**, 253–259 (2010).
73. Majmudar, M. D. *et al.* Monocyte-directed RNAi targeting CCR2 improves infarct healing in atherosclerosis-prone mice. *Circulation* **127**, 2038–2046 (2013).
74. Bursill, C. A., Choudhury, R. P., Ali, Z., Greaves, D. R. & Channon, K. M. Broad-Spectrum CC-Chemokine Blockade by Gene Transfer Inhibits Macrophage Recruitment and Atherosclerotic Plaque Formation in Apolipoprotein E-Knockout Mice. *Circulation* **110**, 2460–2466 (2004).
75. Bursill, C. A. *et al.* Lentiviral gene transfer to reduce atherosclerosis progression by long-term CC-chemokine inhibition. *Gene Ther.* **16**, 93–102 (2009).
76. Weber, C. & Noels, H. Atherosclerosis: current pathogenesis and therapeutic options. *Nat. Med.* **17**, 1410–1422 (2011).
77. Koenen, R. R. & Weber, C. Therapeutic targeting of chemokine interactions in atherosclerosis. *Nat. Rev. Drug Discov.* **9**, 141–153 (2010).

78. McLaren, J. E., Michael, D. R., Ashlin, T. G. & Ramji, D. P. Cytokines, macrophage lipid metabolism and foam cells: implications for cardiovascular disease therapy. *Prog. Lipid Res.* **50**, 331–347 (2011).
79. Wolfs, I. M. J., Donners, M. M. P. C. & de Winther, M. P. J. Differentiation factors and cytokines in the atherosclerotic plaque micro-environment as a trigger for macrophage polarisation. *Thromb. Haemost.* **106**, 763–771 (2011).
80. Chinetti-Gbaguidi, G., Colin, S. & Staels, B. Macrophage subsets in atherosclerosis. *Nat. Rev. Cardiol.* **12**, 10–17 (2015).
81. Leitinger, N. & Schulman, I. G. Phenotypic polarization of macrophages in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **33**, 1120–1126 (2013).
82. Longo, G. M. *et al.* Matrix metalloproteinases 2 and 9 work in concert to produce aortic aneurysms. *J. Clin. Invest.* **110**, 625–632 (2002).
83. Johnson, J. L., George, S. J., Newby, A. C. & Jackson, C. L. Divergent effects of matrix metalloproteinases 3, 7, 9, and 12 on atherosclerotic plaque stability in mouse brachiocephalic arteries. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15575–15580 (2005).
84. Zhao, C. F. & Herrington, D. M. The function of cathepsins B, D, and X in atherosclerosis. *Am. J. Cardiovasc. Dis.* **6**, 163–170 (2016).
85. Liu, J. *et al.* Lysosomal cysteine proteases in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **24**, 1359–1366 (2004).

Chapitre 2 : Vers l'imagerie de la plaque d'athérome

L'athérosclérose, comme vu précédemment, est une maladie qui touche les artères de gros et moyen calibre et cause plus de 19 million de décès par an. Dans la plupart des cas cliniques, la découverte de la plaque ne se fait que suite à un accident vasculaire cérébral ou un infarctus du myocarde. En outre, plus que par l'aspect sténosant de la plaque, la dangerosité est liée à la composition moléculaire et cellulaire. Ainsi, les plaques présentant un large cœur nécrotique recouvert d'une chape fibreuse fine montrent un risque plus important à se rompre et à engendrer la formation d'un thrombus que des plaques considérées comme stables, c'est-à-dire avec une chape fibreuse épaisse et sténosante. Le paramètre le plus caractéristique utilisé en clinique aujourd'hui est la TCFA (Thin-Cap Fibroatheroma). Bien que les méthodes invasives restent les méthodes privilégiées pour le diagnostic de l'athérosclérose en clinique, il existe un intérêt croissant à développer des outils de diagnostic pouvant être utilisés en imagerie non invasive et permettant d'accéder à la composition de l'athérome.

I Imagerie invasive de la plaque

I.1 L'angiographie

Parmi les méthodes d'imagerie anatomique invasive les plus utilisées à l'heure actuelle, l'angiographie occupe le premier rang. Cette technique nécessite l'utilisation d'un cathéter introduit dans l'artère fémorale, pour remonter jusqu'au cœur ainsi que l'usage d'un produit de contraste visible aux rayons X, tel que l'iode. Cette méthode ne permet pas d'atteindre de fortes résolutions et une pénétration suffisante mais reste néanmoins la méthode la plus utilisée en chirurgie interventionnelle. Ainsi, par cette technique, le chirurgien se trouve en mesure de déterminer les endroits de rétrécissement du flux et d'y apposer un stent pour rouvrir la lumière du vaisseau afin de rétablir le flux sanguin (**Figure 10**)^{1,2}.

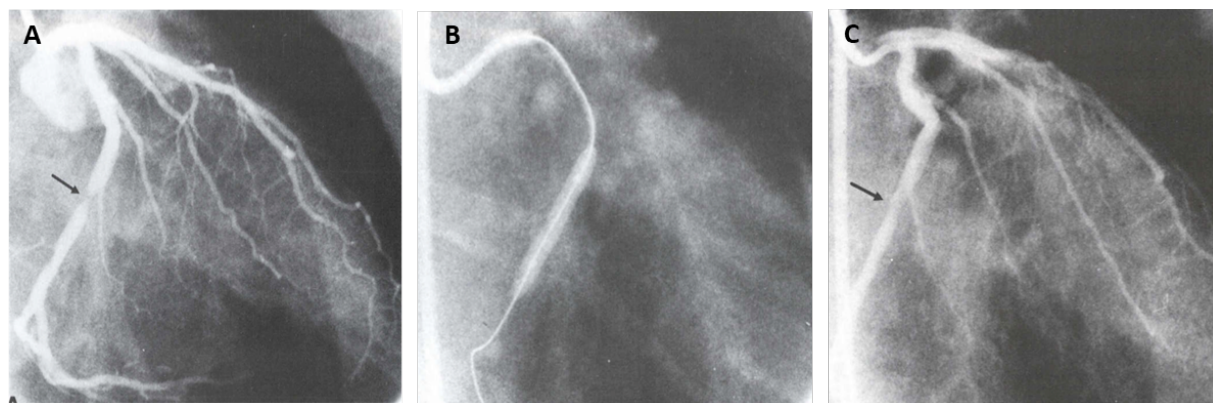


Figure 10 Coronarographie. (A) thrombose de l'artère, (B) mise en place d'un stent, ouverture grâce au gonflement d'un ballonnet, (C) reperfusion de l'artère (Extrait de <http://www.cardio-paramed.com/fr/les-resultats-probleme-restenose.html>)

I.2 La thermographie

Cette technique se base sur la mesure de chaleur au niveau de l'artère. La résolution est de 500 μ m et la précision de 0.05°C³. Elle consiste en l'introduction à l'aide d'un cathéter d'une sonde thermique dans le vaisseau. Ainsi, l'inflammation au niveau de la plaque d'athérome libère de la chaleur qu'il est possible de mesurer grâce au dispositif d'acquisition. En 2003, Naghavi a montré qu'une différence de chaleur pouvait être observée entre un lapin présentant des plaques d'athérome et un lapin sain⁴. Bien que la thermographie renseigne sur l'activité fonctionnelle de la plaque, elle n'apporte pas d'information sur la morphologie de l'artère. Pour contourner cette limitation, l'équipe du Pr Zahi Fayad, en 2001, a proposé de coupler cette méthode de thermographie à l'IVUS (Intravascular Ultra-Sound)⁵.

I.3 L'échographie intravasculaire (IVUS)

L'imagerie IVUS nécessite l'utilisation d'un cathéter avec une sonde ultrason en son extrémité. L'autre extrémité du cathéter est reliée à un système numérique permettant d'acquérir et de visualiser les images. L'IVUS est basée sur la technique des ultrasons avec des transducteurs piézoélectriques qui vont permettre une image de la surface endothéliale interne du vaisseau. Cette technique est la plus fréquemment utilisée dans l'imagerie coronarienne. Elle permet d'évaluer le diamètre du vaisseau, caractériser la plaque et visualiser la paroi artérielle. Les ondes émises se situent entre 20 et 60MHz. Les lipides étant hypoéchogènes et les tissus sains hyperéchogènes, il est possible de différencier une artère saine d'une artère pathologique ^{5,6} (**Figure 11**). Cependant l'échelle de gris IVUS ayant une capacité limitée pour établir la composition de la plaque, une autre technique dérivée de l'IVUS a émergé : l'imagerie IVUS avec une étude en radiofréquence rétrodiffusée nommée VH-IVUS (Vascular Histology – IVUS) ⁷. Cette méthode permet de différencier quatre éléments : la calcification, la fibrose, le cœur nécrotique et la présence de lipides. Du à la présence de nombreux artéfacts et une faible résolution (environ 150 μ m), l'IVUS représente la technique la moins reproductible et fiable pour l'évaluation de la TCFA ⁸, paramètre utilisé à l'heure actuelle en clinique pour le diagnostic des plaques vulnérables ^{9,10}.

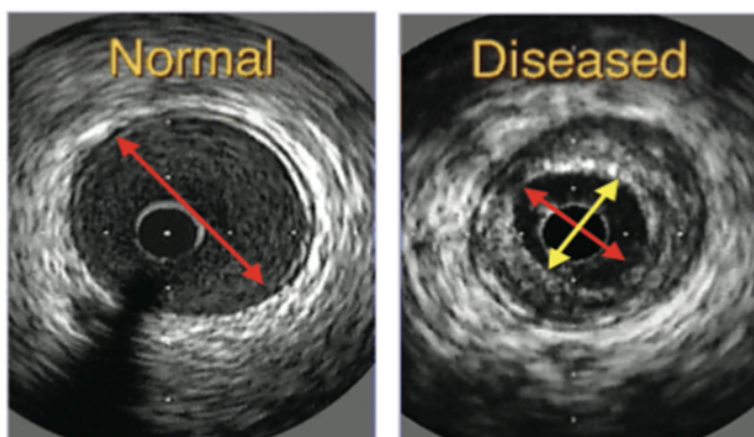


Figure 11 Images IVUS délimitant la lumière d'un vaisseau sain et d'un vaisseau pathologique (Extrait de l'article de Morton Kern paru en 2010 dans *Cath Lab Digest*.)

I.4 La tomographie par cohérence optique (OCT)

Contrairement à l'IVUS qui utilise les ultrasons, la capture d'image en tomographie par cohérence optique s'effectue grâce aux ondes lumineuses en particulier dans l'infra-rouge. La technique est basée sur la notion d'interférométrie à faible cohérence. La résolution de cette technique est de l'ordre du micromètre. L'intérêt de la technologie pour le diagnostic de la plaque d'athérome a été démontré par l'équipe de Fujimoto au MIT (Massachusetts Institute of technology) en 1996 ¹¹, équipe qui avait déjà exposé le principe de l'OCT en 1991 ¹². Une haute résolution et un fort contraste des différents constituants des tissus permettent d'obtenir de façon détaillée les structures et d'avoir la capacité de pénétrer les tissus calcifiés. L'équipe de Otsuka a permis d'imager différents aspects de la plaque : la calcification, la fibrose, l'infiltration de macrophages, la rupture de plaque et le thrombus (**Figure 12**) ¹³. Cette méthode est donc la technique de choix pour imager la TCFA. De ce fait, l'OCT conjugué au développement d'outils de contraste et de ciblage pourrait dans le futur proche être la technique la plus développée pour l'imagerie de la plaque.

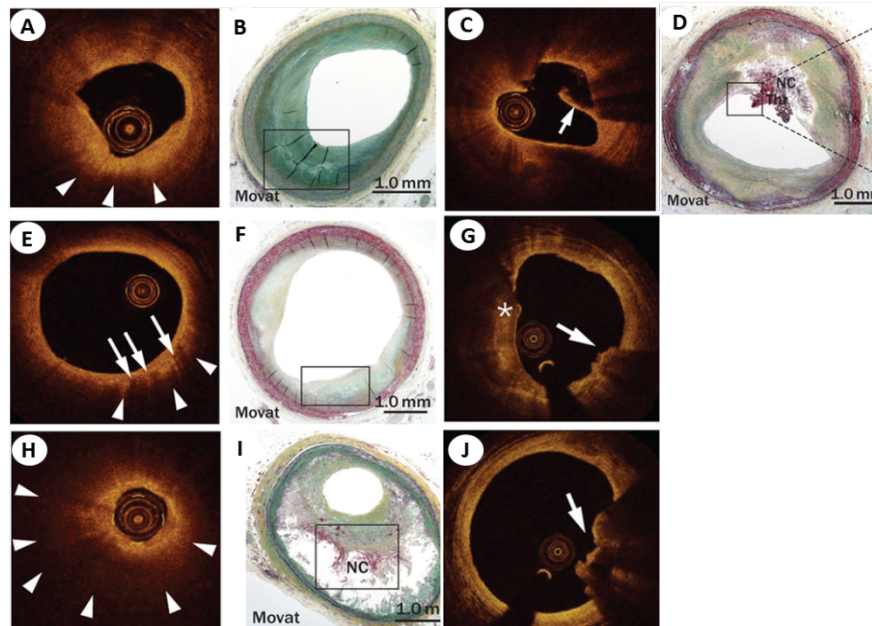


Figure 12 Imagerie OCT des différents aspects de la plaque. (A et B) Images montrant l'épaississement intimal avec peu de macrophages. (C) Les images d'OCT ont révélé la présence d'une rupture de plaque. La rupture de la chape fibreuse montre une région distincte riche en signal. (D) L'histologie a confirmé la présence d'une rupture de la plaque avec un thrombus aigu riche en fibrine. (E et F) Images montrant l'épaississement intimal avec une infiltration de macrophages. La zone avec un fort signal (flèches blanches) atteste de la présence de macrophages spumeux. (A et E) Régions présentant une atténuation du signal focal avec une bordure diffuse (têtes de flèches blanches) dans les images OCT *ex vivo*. (B et F) Présence de lipides en basse résolution. (G et J) Les images OCT montrent la présence de thrombus (flèches blanches) dans les zones de plaque fibrocalcique (astérisque) en l'absence de rupture de la plaque. (H et I) Images montrant un fibroatherome tardif. Les régions pauvres en signal avec une bordure diffuse (têtes de flèche blanches) obtenues en OCT (H) correspondent à la présence de lipides ou d'un corps nécrotique en basse résolution (I) (modifiée à partir de l'article de Otsuka et al. paru en 2014 dans *Nat Rev Cardiol.*)

II Imagerie non invasive de la plaque

II.1 La spectroscopie proche infra-rouge (NIRS)

La spectroscopie proche infra-rouge utilise les longueurs d'ondes entre 780 et 2500nm⁵. Une étude a montré qu'il était possible de corrélérer des données de spectres NIR avec une coloration trichrome histologique de tissus¹⁴. Il a été démontré que la NIRS était capable d'identifier la présence de lipides soulignant donc la présence de plaques vulnérables dans les vaisseaux^{15,16}. Cependant, cette technique n'étant pas en mesure de donner des informations structurales, elle peut être couplée à l'IVUS pour compléter l'analyse (Figure 13)¹⁷.

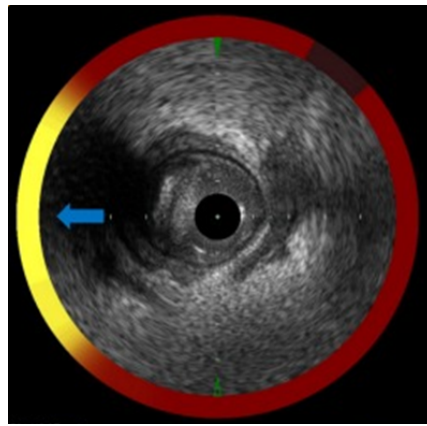


Figure 13 Image co-enregistrée d'IVUS en échelle de gris et NIRS obtenue grâce à un cathéter combiné IVUS-NIRS. La probabilité codifiée par couleur de la présence de lipides, qui est calculée à partir des données de spectres NIRS, est affichée sous forme d'un anneau entourant l'image IVUS en niveaux de gris (flèche bleue) (modifiée à partir de l'article de Yonetsu et al. paru en 2014 dans *Catheter Cardiovasc Interv.*)

II.2 La fluorescence proche infra-rouge (NIRF)

L'imagerie optique utilisant la fluorescence proche infra-rouge est une nouvelle modalité d'imagerie qui a récemment émergé dans le domaine de l'imagerie du cancer. Cette technique possède une forte sensibilité et peut être utilisée pour identifier, localiser et quantifier *in vivo* des sondes moléculaires fluorescentes. L'utilisation de la fluorescence proche infra-rouge, avec des longueurs d'onde comprises entre 700 et 1000nm, permet de minimiser l'auto-fluorescence des tissus mais également une absorption faible des photons qui permet une pénétration plus en profondeur. Afin d'obtenir une image plus résolutive et anatomique, la fluorescence proche infra-rouge a été couplée à l'OCT. Ainsi, elle a permis de donner le profil lipidique de plaque d'athérome (**Figure 14**)¹⁸.

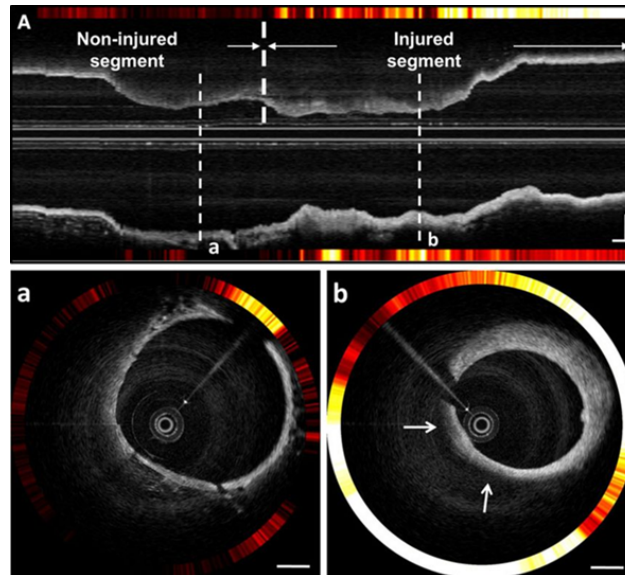


Figure 14 Images de tomographie par cohérence optique (OCT) couplée à la fluorescence proche infrarouge (NIRF) de l'artère de lapin athéroscléreuse *in vivo*. (A) image OCT / NIRF longitudinale de l'artère de lapin injectée par l'indocyanine verte (ICG) (a) Une image OCT / NIRF transversale d'une zone d'apparence normale dans le segment de l'artère d'un lapin seulement sous régime riche en cholestérol (b). Une image OCT / NIRF transversale de la plaque d'athérome dans l'artère de lapin endommagée (avec chirurgie) et sous régime riche en cholestérol. On observe des signaux NIRF importants dans la plaque riche en lipides identifiée par OCT (flèche blanche) (Extrait de l'article de Lee et al. paru en 2014. dans *Circulation : Cardiovascular interventions*)

II.3 La tomodensitométrie (CT = Computed tomography)

La tomodensitométrie a été introduite en pratique clinique en 1972 ^{19,20,21}. C'est une méthode radiographique où un faisceau de rayons X est émis par une source et recueilli par un détecteur placé du côté opposé de la source de rayons X. Le faisceau tourne autour de la zone pour un balayage, tandis que le corps se déplace à travers la machine pour acquérir des données dans les trois directions. L'atténuation du faisceau de rayons X lorsqu'il se déplace à travers les tissus se fait en fonction de la densité des tissus. La reconstruction de ces données collectées en 3D est utilisée pour produire des images tomographiques, images qui sont des "coupes" virtuelles rendant compte du profil d'atténuation du signal de la zone balayée. Pour résumer, la tomodensitométrie est utilisée pour générer une image 3D à partir d'une grande série d'images radiographiques 2D prises autour d'un seul axe de rotation et montrant un contraste basé sur des différences de densité des tissus. La résolution spatiale du CT est de l'ordre de 50 à 200µm.

II.1 Tomographie par émission de positons (TEP)

La TEP est basée sur la détection de deux radiations gamma émises dans deux directions opposées par un atome radioactif après annihilation. Sa résolution spatiale est de 1-2mm. Un traceur radioactif dont la biodistribution et les propriétés biologiques sont connues est injecté. La concentration déterminée en un point précis d'un traceur va donner une image avec échelle de couleur. Le traceur ¹⁸F-FDG (¹⁸F-fluorodesoxyglucose) est le traceur le plus utilisé à l'heure actuelle en imagerie clinique. Ce dernier est métabolisé par la voie du glucose dans les cellules. Il est donc un marqueur de l'activité métabolique. Dans le cadre de l'athérosclérose, le ¹⁸F-FDG est caractéristique de l'augmentation de l'activité macrophagique dans la plaque (**Figure 15**) ^{22,23}.

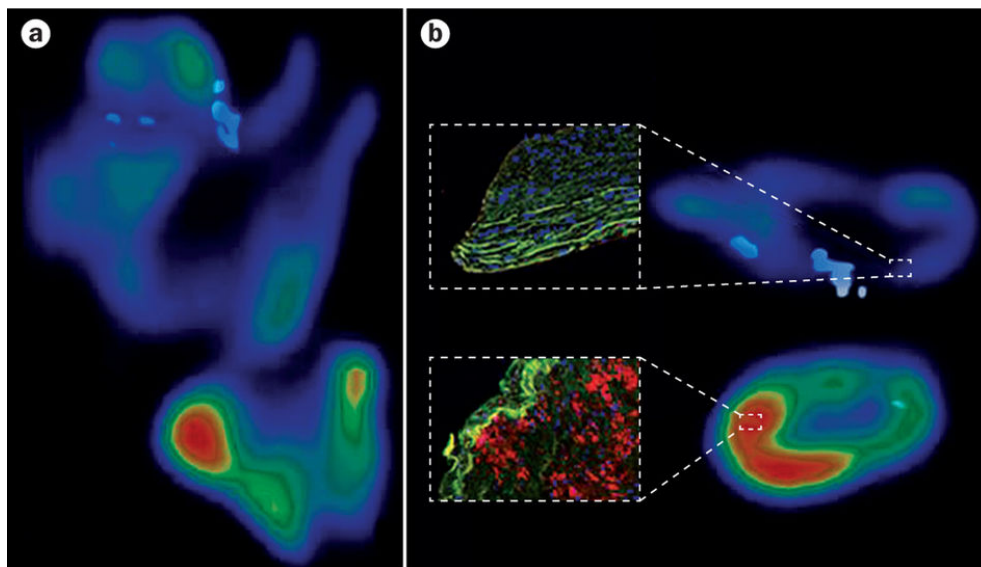


Figure 15 Imagerie de la plaque carotidienne par la technique microPET avec du ¹⁸F-fluorodeoxyglucose (FDG) *ex vivo*. (a) Plan coronal. (b) Plan transversal, avec histologie correspondante montrant une coloration abondante des macrophages CD68+ en corrélation avec les zones d'augmentation de l'absorption de ¹⁸F-FDG au niveau de l'artère carotide (Extrait de l'article de Tarkin et al. paru en 2014 dans *Nat Rev Cardiol.*)

II.2 Tomographie par émission monophotonique (SPECT)

La SPECT permet de reconstruire des images en trois dimensions (3D) des organes et de leur métabolisme à l'aide de caméras gamma qui tournent autour du patient et d'un traceur radioactif. Les caméras gamma sont constituées de collimateurs pour la sélection de photons et de détecteur de rayon gamma. Le spectre des radiotraceurs disponibles pour l'imagerie SPECT est beaucoup plus large que celui de traceurs utilisés en TEP. La résolution spatiale de

la SPECT est identique à celle de la TEP (1-2mm). Parmi les cibles expérimentales de la SPECT, on répertorie les récepteurs IL-2 sur les lymphocytes T²⁴, les récepteurs folates exprimés par les macrophages de type M2 (**Figure 16**)²⁵ ou encore la molécule d'adhésion VCAM exprimée par les cellules endothéliales²⁶.

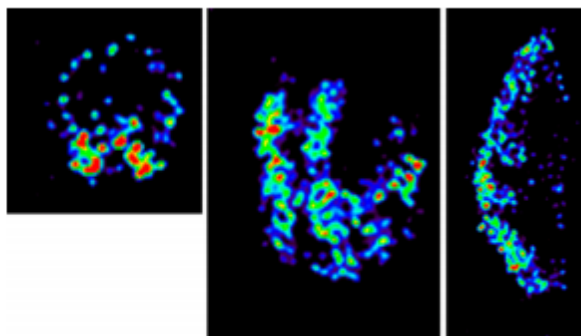


Figure 16 Sections transversale (gauche), coronale (milieu) et sagittale (droite) de l'image micro-SPECT d'un échantillon de plaque carotidienne athéroscléreuse après une incubation avec ^{99m}Tc-folate montrant une accumulation hétérogène. L'image transversale montre que la zone d'accumulation (rouge) est située à l'intérieur de la paroi du vaisseau (Extrait de l'article de Jager et al. paru en 2014 dans *J Nucl Med.*)

II.3 Imagerie par résonance magnétique (IRM)

L'IRM se base sur les propriétés des noyaux des atomes comme le proton de l'eau, qui se comportent comme des petits aimants (appelés « spins ») qui, lorsqu'ils sont placés dans un champ magnétique B_0 , absorbent et réémettent des ondes électromagnétiques à une fréquence proportionnelle à B_0 , appelée fréquence de Larmor. L'IRM consiste à cartographier dans l'espace la résonance des noyaux. Due à leur abondance naturelle, les protons de l'eau et les acides gras composent la majorité du signal. La résolution de l'IRM est de l'ordre de 25 à 100 μ m.

Brièvement, sous l'influence du champ magnétique B_0 (habituellement de 1.5T ou 3T dans les scanners cliniques), les spins sont excités grâce à des radiofréquences (ondes électromagnétiques) produites par une antenne d'émission. Après excitation, les spins vont retourner à leur état d'équilibre, appelé « relaxation » en émettant des radiofréquences qui sont récoltées par l'antenne d'émission, créant ainsi le signal RM (**Figure 17**). Pour construire une image, le signal RM est encodé dans l'espace en phase et en fréquence grâce à un gradient magnétique (30 à 45 mT/m) superposé avant et pendant la réémission des radiofréquences. En IRM, le signal mesuré est déterminé par la séquence d'impulsion qui dépend de (1) la densité des spins nucléaires associée aux atomes d'hydrogènes, appelée « densité de proton », (2) du temps de relaxation T1 spin-réseau qui reflète le temps que la magnétisation longitudinal prend pour retourner à l'état d'équilibre et (3) du temps de relaxation spin-spin, aussi connu comme le temps de relaxation T2, qui représente la perte de cohérence de l'aimantation transversale. Ces propriétés sont intrinsèques aux tissus, ce qui permet, en adaptant la séquence d'impulsion, de créer un contraste tissulaire RM, telles que les images pondérées T1 ou T2.

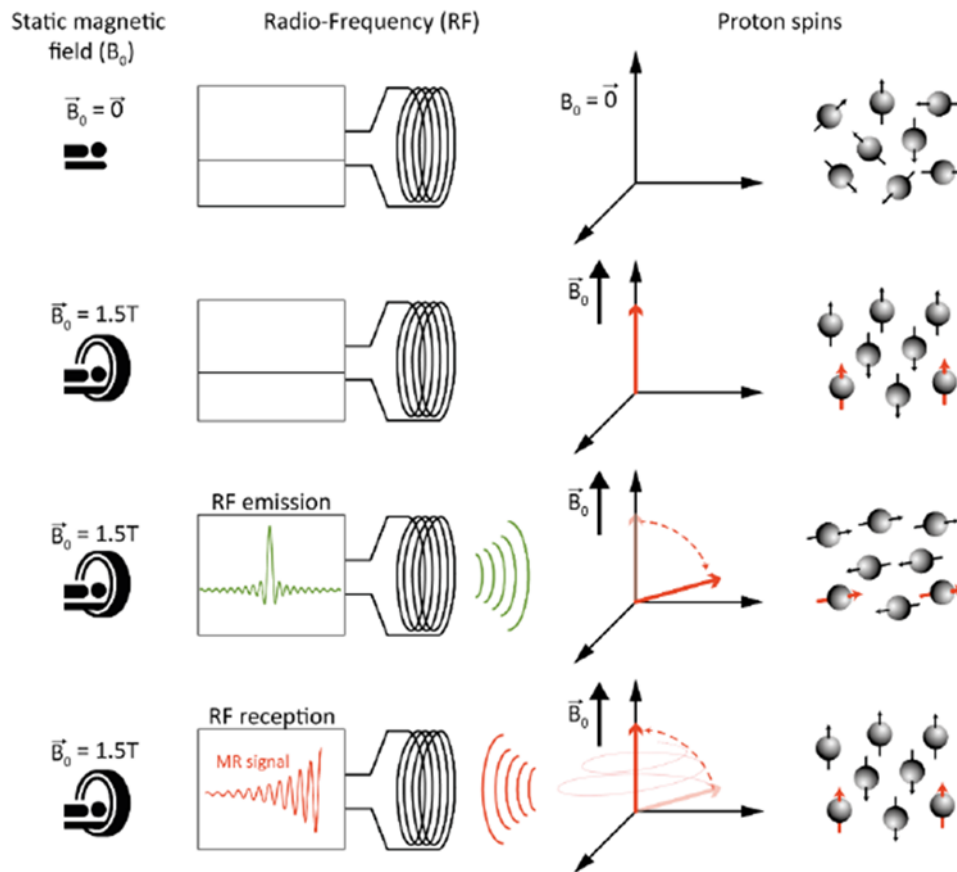


Figure 17 Principe de l'IRM. La magnétisation \mathbf{M} , peut être décomposée en deux vecteurs : l'aimantation longitudinale, selon \mathbf{B}_0 et l'aimantation transversale, dans le plan \mathbf{Oxy} . Lors de l'impulsion radio-fréquence, à la fréquence de Larmor, l'aimantation bascule dans le plan transversal, puis revient à son état d'équilibre en suivant un mouvement hélicoïdal. Ce retour à l'état d'équilibre est appelé « relaxation ». Les spins qui reviennent à l'état d'équilibre émettent des radiofréquences qui constituent le signal RM (Schématisation de l'IRM par Cyril Lorenzato – membre post-doc de l'équipe de Gisèle CLOFENT-SANCHEZ) (à écrire –détails).

L'IRM présente un fort potentiel dans le phénotypage des maladies vasculaires car il pourrait être utilisé pour différencier les composants de la plaque sur la base de propriétés biophysiques et biochimiques telles que la composition chimique, la teneur en eau, l'état physique, le mouvement moléculaire ou la diffusion²⁷. La plus grande force de l'IRM pour caractériser la plaque d'athérosclérose est la disponibilité de protocoles multi-contrastes pondérés T1-T2 utilisant les techniques du « sang clair », du « sang noir » et du temps de vol (TOF). Tandis que les séquences « sang clair » n'affectent pas le signal sanguin, les séquences « sang noir » comptent sur l'élimination du signal du flux sanguin et représentent une approche générale pour la caractérisation de la paroi du vaisseau. De manière opposée, les séquences TOF ont tendance à minimiser le signal de tissus stationnaires permettant la visualisation de la lumière vasculaire. Par exemple, l'équipe de Chu a développé une méthode de cartographie, appelée Morphology Enhanced Probability maPS (MEPPS), permettant de distinguer différents composants de la plaque grâce à la combinaison d'images TOF et de sang noir T1 / images pondérées T1 / T2 renforcées par le contraste (CE) (**Figure 18**). L'injection du produit de contraste, le gadolinium, a permis de créer une différence de contraste au niveau de la plaque carotidienne. Cette forte différence de signal est souvent associée à une perméabilité du tissu en un endroit donné qui signe une absence de la matrice extracellulaire.

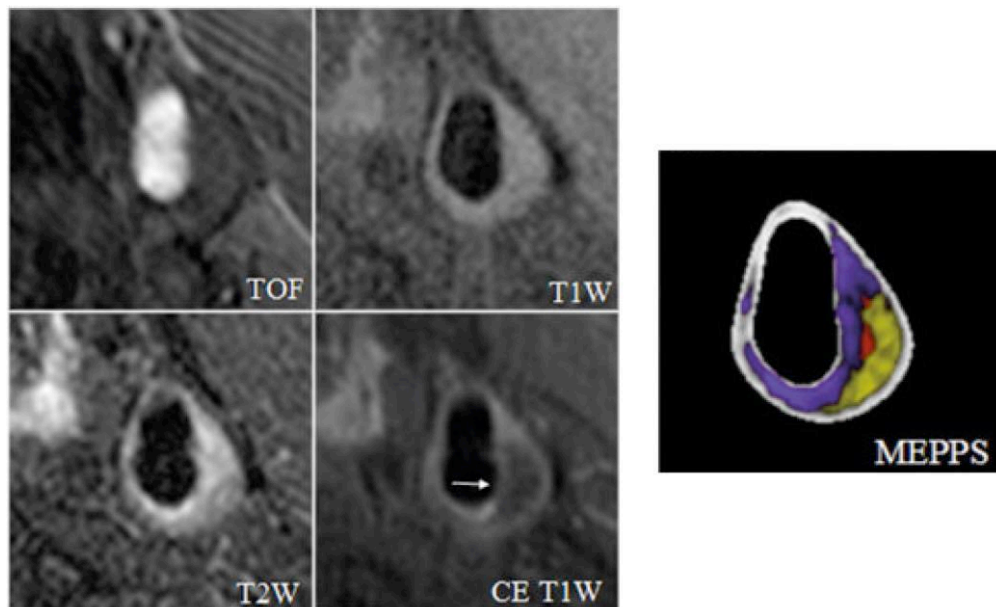


Figure 18 Caractérisation structurale des plaques d'athérosclérose par une combinaison de séquences. Temps de vol (TOF), pondéré en T1 (T1W), pondéré en T2 (T2W) et contraste renforcé (par injection d'un agent de contraste : le gadolinium) MR pondéré en T1 (CE-T1W). Les images sont combinées pour produire la Morphology Enhanced Probability maPS (MEPPS). La perte de matrice est montrée dans le noyau nécrotique (violet), ainsi que dans la zone riche en lipides (jaune), et l'hémorragie intraplaque (rouge) sur l'image MEPPS (Extrait de l'article de Chu et al. paru en 2009 dans la revue *JACC*)

Comme beaucoup de méthodes d'imagerie citées ci-dessus, la combinaison de l'IRM est également envisagée notamment avec les méthodes comme le PET afin d'augmenter la sensibilité. L'IRM permet d'apporter l'aspect anatomique alors que le PET permet d'apporter une meilleure sensibilité et une meilleure résolution. Une étude pré-clinique récente a montré l'avantage que pouvait avoir ce couplage en comparaison à l'IRM seul et au PET/CT ²⁸.

La dangerosité de la plaque n'étant pas seulement liée à sa structure mais également à la composition cellulaire et moléculaire, les équipes de recherche ont tenté de développer des objets capables de cibler la plaque. Ainsi, si l'on prend en exemple le rôle clé des macrophages dans l'inflammation et notamment leur activité phagocytaire, Briley-Saebo et al. ont montré que des USPIO (Ultrasmall superparamagnetic iron oxide) recouverts de dextran (Feridex) avec une taille inférieure à 20nm s'accumulaient dans les macrophages intraplaques d'un lapin athéromateux ²⁹. Outre le ciblage passif des macrophages, d'autres agents ciblant cette même population de cellules ont été développés à partir de ligands physiologiques pour apporter, dans un premier temps, un agent de diagnostic. Des micelles couplées à des anticorps anti LDL oxydées, en particulier IK17, anticorps découvert par Peter Shaw par phage display *in vitro* et chargées en gadolinium ont montré une fixation spécifique au niveau de la plaque d'athérome dès 24h jusqu'à 72h après injection chez des souris ApoE^{-/-}. L'imagerie a été réalisée par IRM (**Figure 19**)³⁰.

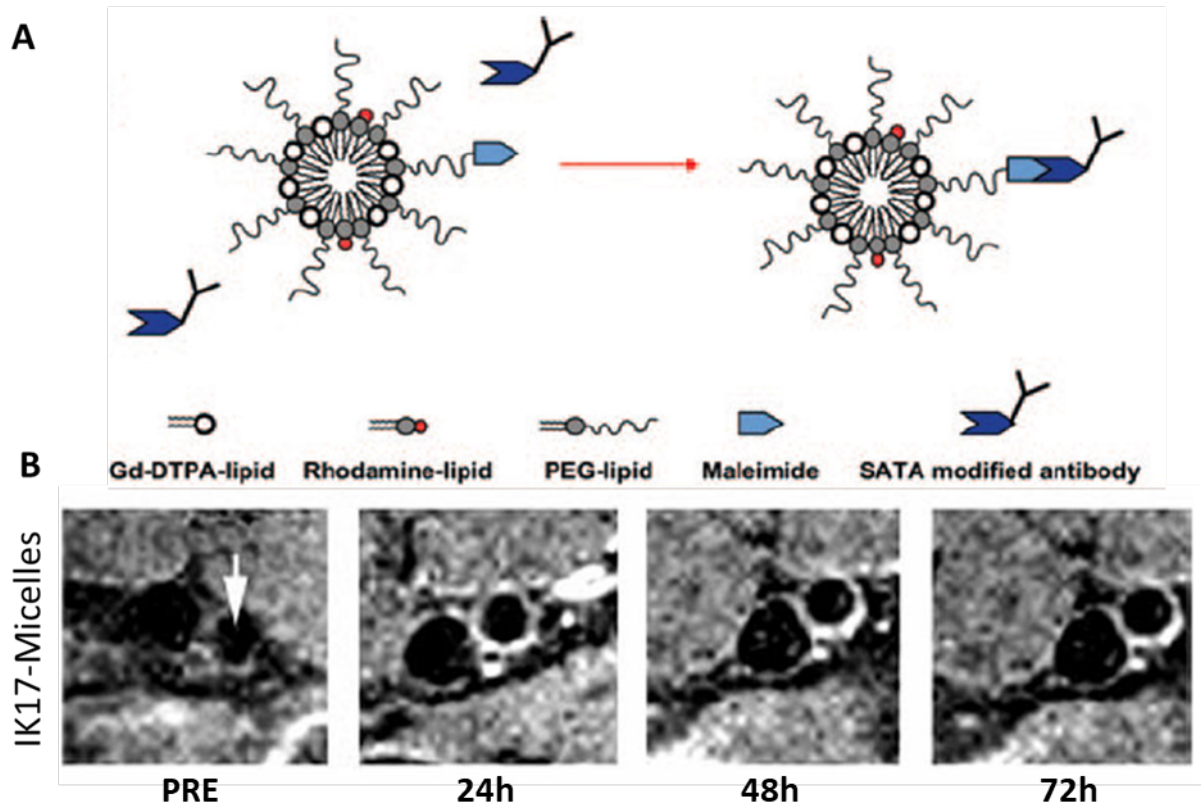


Figure 19 (A) Schéma de l'attachement de l'S-acéthylthioglycolique acid N-hydroxysuccinimide ester (SATA) aux anticorps, de la composition des micelles et de la fixation des anticorps aux micelles. (B) Imagerie de l'aorte abdominale (flèches) chez les souris apoE - / - en fonction du temps après injection de 0,075 mmol Gd/ kg de micelles-Anticorps. Le vaisseau à gauche de l'aorte est la veine cave inférieure (IVC). Bien que le flux artériel ait été saturé pour permettre la délimitation de la paroi artérielle, un bruit de fond au niveau de l'IVC peut être observé dû à l'artéfact d'écoulement (Modifié à partir de l'article de Briley-Saebo et al. paru en 2015)

Il a également été développé des agents ciblant apportant à la fois l'agent de diagnostic mais également un agent de thérapie. En effet, une autre étude a montré que des nanoparticules « lipoprotéine de haute densité reconstituée » (rHDL) chargées en statines mais également d'un agent de contraste, le gadolinium, vont, grâce à l'ApoA1, faciliter la délivrance de statines au niveau de la plaque d'athérome. La biodistribution des rHDL a été suivie pendant 12 semaines à l'aide d'images RM pondérées en T1 des aortes abdominales de souris (**Figure 20b**)³¹. L'accumulation des rHDL marquées au cyanine 5.5 au niveau des plaques d'athérome a été également montrée par image NIRF après injection par voie intraveineuse à des souris ApoE KO (**Figure 20c**).

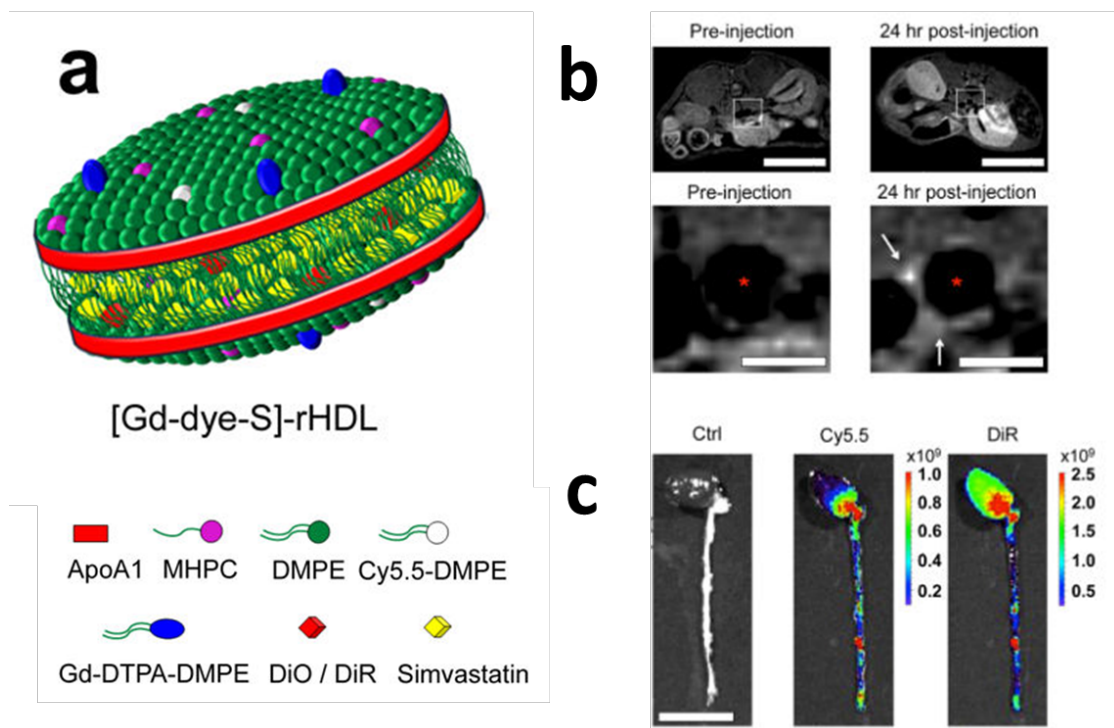


Figure 20 (a) Représentation schématique d'une lipoprotéine de haute densité reconstituée contenant des statines ([Gd-dye-S] -rHDL), contenant du gadolinium et un fluorochrome (Cy5.5, DiO, DiR); (b) Images pondérées en T1 de l'aorte abdominale d'une souris ApoE ^{-/-} avant et 24 heures après l'injection de [Gd-Dye-S] -rHDL. La lumière est indiquée par une (*). L'image de post-injection à 24h montre l'amélioration du signal dans la paroi du vaisseau (flèches blanches), indiquant une infiltration et une rétention des nanoparticules dans les plaques aortiques. (c) Des [S]-rHDL marquées avec le Cy5.5 (monocouche lipidique) et le DiR (noyau hydrophobe) ont été injectées par voie intraveineuse à des souris ApoE ^{-/-}. L'image NIRF montre que Cy5.5 et DiR s'accumulent préférentiellement dans les zones riches en lipides au niveau des lésions athéroscléreuse (modifiée à partir de l'article de Diuvenvoorden et al. paru en 2014 dans la revue *Nat Comm*)

Dans cette étude, les lipoprotéines ApoA1, qui structurent la rHDL (voir la **Figure 20a**), ont été utilisées comme agents de ciblage pour délivrer les statines au site d'intérêt. Cependant, l'ApoA1 est non spécifique aux macrophages colonisant la plaque d'athérome. Ainsi, la clé pour contourner l'accumulation hépatique et l'hépatotoxicité repose sur la fonctionnalisation des HDL avec des ligands spécifiques, tels que les anticorps, afin de cibler de manière spécifique la plaque d'athérome.

III Conclusion

L'athérosclérose, maladie chronique et inflammatoire, représente aujourd'hui un réel fléau économique et social pour la société dans les pays développés. De nos jours, il existe donc un intérêt croissant pour la détection des plaques à caractère instable considérées comme étant plus susceptibles de se rompre que des plaques sténosantes. Les recherches se sont ainsi tournées ces dernières années vers le développement d'outils de diagnostic capables de cibler des facteurs impliqués dans la progression de la formation des plaques (**Figure 21**). La première étape de développement de la plaque étant l'activation de l'endothélium, des outils ont été développés pour cibler les molécules d'adhésion VCAM-1 ^{32,33,34}, la P-sélectine ³⁵ ou encore la E-sélectine ³⁶. Après activation de l'endothélium, des cellules immunes vont être recrutées. Impliquée et responsable de la diapédèse, la molécule MCP-1 fait également partie des cibles de l'imagerie car elle est le premier marqueur inflammatoire dans l'athérogénèse ^{37,38}. Impliqués dans le processus évolutif et marqueurs de l'instabilité de la plaque, les macrophages et leur activité métabolique font aujourd'hui l'objet de recherche de plusieurs équipes. Le ¹⁸F-FDG, métabolisé dans la glycolyse au niveau du macrophage, a été utilisé

pour l'imagerie TEP^{39,22,23}. Le stress oxydatif, lié à l'activité macrophagique, est également la cible de sondes fluorescentes^{40,41,42}. Parmi les facteurs influençant la stabilité de la plaque se trouvent les métalloprotéinases sécrétées par les macrophages intra-plaques. L'équipe de Pr Zahi A. Fayad a mis au point des sondes peptidiques pour l'IRM qui ciblent ces métalloprotéinases⁴³.

Un autre processus qui accompagne le développement de la plaque, l'angiogénèse, est caractérisé par l'expression de molécules comme les intégrine $\alpha\beta_3$, exprimées au niveau des néovaisseaux. Parmi les équipes de recherche qui s'intéressent à ces molécules, Laitinen a développé un peptide dirigé contre ces molécules pour une imagerie TEP/CT⁴⁴. Des sondes ciblant les LDL oxydées au niveau du cœur nécrotique mais également le thrombus lié à la rupture de ce cœur lipidique ont également été développés durant ces 20 dernières années pour une imagerie SPECT mais également IRM^{45,46,30}. Sous l'action d'enzymes de dégradation mais également de molécules issues du stress oxydatif, les cellules macrophagiques constituant le cœur nécrotique rentrent en apoptose. Des sondes ont donc été développées dans le but de cibler les cellules apoptotiques, et plus précisément la phosphatidylsérine^{47,48}. Jusqu'à aujourd'hui, afin de diagnostiquer la plaque et ces différents stades de progression, les chercheurs ont utilisé et développé des peptides ou bien des anticorps murins destinés à la reconnaissance de cibles murines (modèle pré-clinique le plus utilisé en recherche). Cependant, deux des problèmes majeurs liés à l'utilisation d'anticorps de rongeurs ou de peptides sont d'une part l'affinité vis-à-vis de cibles thérapeutiques humaines mais également l'immunogénicité une fois injectés chez l'homme. Pour pallier cela, l'humanisation des anticorps est une étape nécessaire. Par la suite, Peter Shaw a ouvert la voie à de nouvelles perspectives : le développement d'anticorps entièrement humains présentant l'avantage d'une meilleure affinité mais également d'une immunogénicité moindre. En effet, un anticorps anti LDL oxydées, nommé IK17, a été isolé par phage display *in vitro* et a montré après couplage à des micelles chargées en gadolinium un ciblage spécifique de la plaque d'athérome⁴⁹. Ainsi, ces anticorps étant entièrement humains pourront être transférés directement dans les études cliniques.

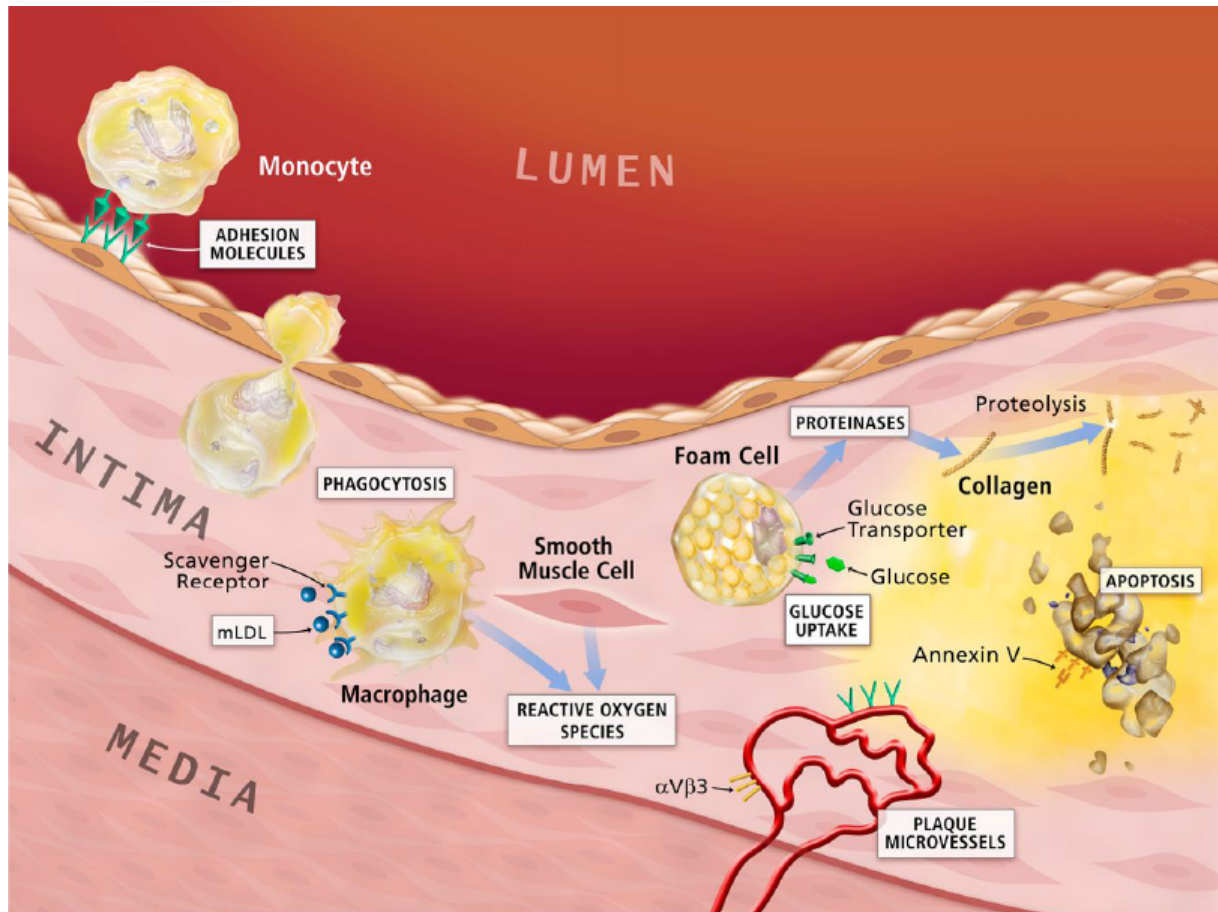


Figure 21 Les cibles d'imagerie moléculaire potentielles dans l'athérosclérose. L'athérogenèse implique le recrutement de cellules inflammatoires à partir du sang, représenté par le monocyte dans le coin supérieur gauche de ce diagramme. Les monocytes sont les leucocytes les plus nombreux dans la plaque athérosclérotique. Le recrutement dépend de l'expression des molécules d'adhésion sur l'endothélium macrovasculaire, comme indiqué, et sur les microvaisseaux de plaque. Une fois résidant dans l'intima artérielle, les macrophages activés deviennent phagocytiquement actifs, un processus qui fournit une autre cible potentielle pour l'imagerie de la plaque. Des epitopes associés à des lipoprotéines de faible densité (LDL) oxydées qui s'accumulent dans des plaques peuvent également servir de cibles pour l'imagerie moléculaire. Les cellules spumeuses peuvent présenter une activité métabolique accrue, augmentant leur absorption de glucose, un processus déjà mesurable dans la clinique par l'absorption de ^{18}F -FDG. Les phagocytes activés peuvent également élaborer des enzymes dégradant les protéines qui peuvent cataboliser le collagène dans la chape fibreuse de la plaque, l'affaiblir et la rendre sensible à la rupture et, par conséquent, à la thrombose. Les phagocytes mononucléaires qui meurent par apoptose dans les plaques présentent des niveaux accrus de phosphatidylsérine à leur surface. Des sondes pour l'apoptose telles que l'annexine V peuvent également permettre de visualiser des athéromes compliqués. Les microvaisseaux eux-mêmes peuvent exprimer non seulement les molécules d'adhésion des leucocytes (indiquées en vert) mais aussi les intégrines telles que $\alpha\text{V}\beta\text{3}$. Les encadrés blancs mettent en exergue des voies aux cibles potentielles (Extrait de l'article de Libby et al. paru en 2010 dans la revue *JNM*)

Bibliographie

1. Patel, M. R. *et al.* Low diagnostic yield of elective coronary angiography. *N. Engl. J. Med.* **362**, 886–895 (2010).
2. Marwick, T. H., Cho, I., Ó Hartaigh, B. & Min, J. K. Finding the Gatekeeper to the Cardiac Catheterization Laboratory: Coronary CT Angiography or Stress Testing? *J. Am. Coll. Cardiol.* **65**, 2747–2756 (2015).
3. Stefanadis, C. *et al.* Thermal heterogeneity within human atherosclerotic coronary arteries detected in vivo: A new method of detection by application of a special thermography catheter. *Circulation* **99**, 1965–1971 (1999).
4. Naghavi, M. *et al.* Thermography basket catheter: in vivo measurement of the temperature of atherosclerotic plaques for detection of vulnerable plaques. *Catheter. Cardiovasc. Interv. Off. J. Soc. Card. Angiogr. Interv.* **59**, 52–59 (2003).
5. Fayad, Z. A. & Fuster, V. Clinical imaging of the high-risk or vulnerable atherosclerotic plaque. *Circ. Res.* **89**, 305–316 (2001).
6. Reiber, J. H. C. & Serruys, P. W. *Progress in quantitative coronary arteriography.* (Springer Science & Business Media, 2012).
7. Nair, A., Margolis, M. P., Kuban, B. D. & Vince, D. G. Automated coronary plaque characterisation with intravascular ultrasound backscatter: ex vivo validation. *EuroIntervention J. Eur. Collab. Work. Group Interv. Cardiol. Eur. Soc. Cardiol.* **3**, 113–120 (2007).
8. Garcia-Garcia, H. M. *et al.* Imaging plaques to predict and better manage patients with acute coronary events. *Circ. Res.* **114**, 1904–1917 (2014).
9. Narula, J. *et al.* Histopathologic characteristics of atherosclerotic coronary disease and implications of the findings for the invasive and noninvasive detection of vulnerable plaques. *J. Am. Coll. Cardiol.* **61**, 1041–1051 (2013).

10. Redgrave, J. N. E., Lovett, J. K., Gallagher, P. J. & Rothwell, P. M. Histological assessment of 526 symptomatic carotid plaques in relation to the nature and timing of ischemic symptoms: the Oxford plaque study. *Circulation* **113**, 2320–2328 (2006).
11. Brezinski, M. E. *et al.* Optical coherence tomography for optical biopsy. Properties and demonstration of vascular pathology. *Circulation* **93**, 1206–1213 (1996).
12. Huang, D. *et al.* Optical coherence tomography. *Science* **254**, 1178–1181 (1991).
13. Otsuka, F., Joner, M., Prati, F., Virmani, R. & Narula, J. Clinical classification of plaque morphology in coronary disease. *Nat. Rev. Cardiol.* **11**, 379–389 (2014).
14. Moreno, P. R. *et al.* Detection of lipid pool, thin fibrous cap, and inflammatory cells in human aortic atherosclerotic plaques by near-infrared spectroscopy. *Circulation* **105**, 923–927 (2002).
15. Gardner, C. M. *et al.* Detection of lipid core coronary plaques in autopsy specimens with a novel catheter-based near-infrared spectroscopy system. *JACC Cardiovasc. Imaging* **1**, 638–648 (2008).
16. Waxman, S. *et al.* In vivo validation of a catheter-based near-infrared spectroscopy system for detection of lipid core coronary plaques: initial results of the SPECTACL study. *JACC Cardiovasc. Imaging* **2**, 858–868 (2009).
17. Yonetsu, T. *et al.* Comparison of Near-Infrared Spectroscopy and Optical Coherence Tomography for Detection of Lipid. *Catheter. Cardiovasc. Interv. Off. J. Soc. Card. Angiogr. Interv.* **84**, (2014).
18. Kim, S. *et al.* Intracoronary dual-modal optical coherence tomography-near-infrared fluorescence structural–molecular imaging with a clinical dose of indocyanine green for the assessment of high-risk plaques and stent-associated inflammation in a beating coronary artery. *Eur. Heart J.* **37**, 2833–2844 (2016).
19. Cormack, A. M. Representation of a Function by Its Line Integrals, with Some Radiological Applications. *J. Appl. Phys.* **34**, 2722–2727 (1963).

20. Hounsfield, G. N. Computerized transverse axial scanning (tomography). 1. Description of system. *Br. J. Radiol.* **46**, 1016–1022 (1973).
21. Kalender, W. A. X-ray computed tomography. *Phys. Med. Biol.* **51**, R29-43 (2006).
22. Tarkin, J. M., Joshi, F. R. & Rudd, J. H. F. PET imaging of inflammation in atherosclerosis. *Nat. Rev. Cardiol.* **11**, 443–457 (2014).
23. Ogawa, M., Nakamura, S., Saito, Y., Kosugi, M. & Magata, Y. What can be seen by 18F-FDG PET in atherosclerosis imaging? The effect of foam cell formation on 18F-FDG uptake to macrophages in vitro. *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.* **53**, 55–58 (2012).
24. Glaudemans, A. W. J. M. *et al.* In vivo and in vitro evidence that ^{99m}Tc-HYNIC-interleukin-2 is able to detect T lymphocytes in vulnerable atherosclerotic plaques of the carotid artery. *Eur. J. Nucl. Med. Mol. Imaging* **41**, 1710–1719 (2014).
25. Jager, N. A. *et al.* Folate receptor- β imaging using ^{99m}Tc-folate to explore distribution of polarized macrophage populations in human atherosclerotic plaque. *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.* **55**, 1945–1951 (2014).
26. Broisat, A. *et al.* ^{99m}Tc-cAbVCAM1-5 imaging is a sensitive and reproducible tool for the detection of inflamed atherosclerotic lesions in mice. *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.* **55**, 1678–1684 (2014).
27. Phinikaridou, A. *et al.* Molecular MRI of Atherosclerosis. *Mol. Basel Switz.* **18**, 14042–14069 (2013).
28. Ripa, R. S. & Kjær, A. Imaging Atherosclerosis with Hybrid Positron Emission Tomography/Magnetic Resonance Imaging. *BioMed Res. Int.* **2015**, (2015).
29. Briley-Saebo, K. C., Mani, V., Hyafil, F., Cornily, J.-C. & Fayad, Z. A. Fractionated Feridex and positive contrast: in vivo MR imaging of atherosclerosis. *Magn. Reson. Med.* **59**, 721–730 (2008).
30. Briley-Saebo, K. C. *et al.* Targeted Molecular Probes for Imaging Atherosclerotic Lesions With Magnetic Resonance Using Antibodies That Recognize Oxidation-Specific Epitopes. *Circulation* **117**, 3206–3215 (2008).

31. Duivenvoorden, R. *et al.* A statin-loaded reconstituted high-density lipoprotein nanoparticle inhibits atherosclerotic plaque inflammation. *Nat. Commun.* **5**, 3065 (2014).
32. Nahrendorf, M. *et al.* Noninvasive vascular cell adhesion molecule-1 imaging identifies inflammatory activation of cells in atherosclerosis. *Circulation* **114**, 1504–1511 (2006).
33. Nahrendorf, M. *et al.* ¹⁸F-4V for PET-CT imaging of VCAM-1 expression in inflammatory atherosclerosis. *JACC Cardiovasc. Imaging* **2**, 1213–1222 (2009).
34. McAteer, M. A. *et al.* Magnetic resonance imaging of endothelial adhesion molecules in mouse atherosclerosis using dual-targeted microparticles of iron oxide. *Arterioscler. Thromb. Vasc. Biol.* **28**, 77–83 (2008).
35. Jacobin-Valat, M.-J. *et al.* MRI of inducible P-selectin expression in human activated platelets involved in the early stages of atherosclerosis. *NMR Biomed.* **24**, 413–424 (2011).
36. Kang, H. W., Josephson, L., Petrovsky, A., Weissleder, R. & Bogdanov, A. Magnetic resonance imaging of inducible E-selectin expression in human endothelial cell culture. *Bioconjug. Chem.* **13**, 122–127 (2002).
37. Ohtsuki, K., Hayase, M., Akashi, K., Kapiwoda, S. & Strauss, H. W. Detection of monocyte chemoattractant protein-1 receptor expression in experimental atherosclerotic lesions: an autoradiographic study. *Circulation* **104**, 203–208 (2001).
38. Hartung, D. *et al.* Radiolabeled Monocyte Chemotactic Protein 1 for the detection of inflammation in experimental atherosclerosis. *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.* **48**, 1816–1821 (2007).
39. Rudd, J. H. F. *et al.* Imaging atherosclerotic plaque inflammation with [¹⁸F]-fluorodeoxyglucose positron emission tomography. *Circulation* **105**, 2708–2711 (2002).
40. Li, D. *et al.* Molecular Imaging of Atherosclerotic Plaques Targeted to Oxidized LDL Receptor LOX-1 Using SPECT/CT and Magnetic Resonance. *Circ. Cardiovasc. Imaging* **3**, 464–472 (2010).

41. Amirbekian, V. *et al.* Detecting and assessing macrophages in vivo to evaluate atherosclerosis noninvasively using molecular MRI. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 961–966 (2007).
42. Lipinski, M. J. *et al.* Macrophage-specific lipid-based nanoparticles improve cardiac magnetic resonance detection and characterization of human atherosclerosis. *JACC Cardiovasc. Imaging* **2**, 637–647 (2009).
43. Amirbekian, V. *et al.* Atherosclerosis and matrix metalloproteinases: experimental molecular MR imaging in vivo. *Radiology* **251**, 429–438 (2009).
44. Laitinen, I. *et al.* Evaluation of alphavbeta3 integrin-targeted positron emission tomography tracer 18F-galacto-RGD for imaging of vascular inflammation in atherosclerotic mice. *Circ. Cardiovasc. Imaging* **2**, 331–338 (2009).
45. von zur Muhlen, C. *et al.* Visualization of activated platelets by targeted magnetic resonance imaging utilizing conformation-specific antibodies against glycoprotein IIb/IIIa. *J. Vasc. Res.* **46**, 6–14 (2009).
46. Taillefer, R., Edell, S., Innes, G. & Lister-James, J. Acute thromboscintigraphy with (99m)Tc-apcitide: results of the phase 3 multicenter clinical trial comparing 99mTc-apcitide scintigraphy with contrast venography for imaging acute DVT. Multicenter Trial Investigators. *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.* **41**, 1214–1223 (2000).
47. Burtea, C. *et al.* Peptidic targeting of phosphatidylserine for the MRI detection of apoptosis in atherosclerotic plaques. *Mol. Pharm.* **6**, 1903–1919 (2009).
48. Kietzelaer, B. L. J. H. *et al.* Noninvasive Detection of Plaque Instability with Use of Radiolabeled Annexin A5 in Patients with Carotid-Artery Atherosclerosis. *N. Engl. J. Med.* **350**, 1472–1473 (2004).
49. Shaw, P. X. *et al.* Human-Derived Anti-Oxidized LDL Autoantibody Blocks Uptake of Oxidized LDL by Macrophages and Localizes to Atherosclerotic Lesions In Vivo. *Arterioscler. Thromb. Vasc. Biol.* **21**, 1333–1339 (2001).

Chapitre 3: Les anticorps

I Histoire

En 1798, Edward Jenner propose l'inoculation d'une maladie bénigne pour l'homme, la vaccine ou variole des bovidés pour le protéger contre une affection humaine grave, la variole¹. C'est en 1880 que Louis Pasteur montre que les Gallinacées peuvent devenir résistants au choléra des poules en leur injectant une vieille culture du germe dans laquelle l'agent pathogène a perdu une partie de sa virulence². Cette méthode sera nommée par ce dernier, vaccination, terme choisi en l'hommage aux travaux de Jenner sur la vaccine. En 1885, Pasteur développa la première vaccination antirabique par germes inactivés. Parallèlement aux études de Pasteur, en 1882, Elie Metchnikoff découvre la phagocytose d'agents bactériens par les macrophages, constituant un des volets de l'immunité dite « innée », qui est non spécifique³.

En 1888, Emile Roux et Alexandre Yersin démontrèrent que le transfert de sérum d'un animal immunisé rendait l'animal sain résistant contre la toxine diphtérique (d'après « *Une histoire des microbes* » de P. Berche, p.216-217). En 1890, Emil Von Behring et Shibasaburo Kitasato firent la même constatation en montrant que des cobayes immunisés contre l'anatoxine diphtérique devenaient résistants à l'administration ultérieure d'une dose létale. Cependant, ils découvrirent également que ces animaux immunisés contre la toxine diphtérique succombaient à la toxine tétanique, émettant alors la notion de spécificité⁴. Ces auteurs montrèrent alors que les facteurs responsables de cette immunité étaient contenus dans le sérum des animaux vaccinés. Les substances capables de neutraliser ces toxines furent alors appelées anti-toxines. C'est en 1897 que Paul Ehrlich donna le nom d'anticorps aux « anti-toxines » définissant en retour les antigènes comme molécules reconnues et neutralisées par les anticorps. C'est ainsi que naissait la sérothérapie, initiant une révolution médicale, celle des anticorps thérapeutiques (d'après « *Immunologie – 6^e édition* » de Chatenoud et Bach). Entre 1895 et 1899, Jules Bordet montra que deux facteurs existants dans le sérum permettaient la destruction des bactéries appelée « bactériolyse ». Ces deux éléments sont (1) des anticorps thermostables se fixant de manière spécifique à ces bactéries et présents dans le sérum d'organismes déjà immunisés; (2) un facteur thermosensible présent chez tous les animaux qu'il nomma à l'époque alexine et nommé aujourd'hui « complément ». Trois années plus tard, il mettra également en évidence que les globules rouges d'une autre espèce dans un sérum sont lysés par l'action du complément également⁵.

En 1948, Astrid Fagreus a montré que les anticorps étaient produits par des plasmocytes, un type cellulaire dérivé de l'activation des lymphocytes B (d'après « *Handbook of experimental immunology* » de D.M. Weir). Franck MacFarlane Burnet, médecin australien, introduit, en 1957, la notion de sélection clonale qui montre qu'un lymphocyte B n'exprime à sa surface qu'un seul type de récepteur spécifique d'un seul épitope. Ce récepteur membranaire est un anticorps nommé BCR (B-cell receptor)⁶. En 1960, une autre catégorie de lymphocytes fut identifiée, les lymphocytes T exprimant à leur surface des récepteurs membranaires appelés TCR (T Cell Receptor).

Une deuxième théorie vit alors le jour, celle de l'immunité adaptative qui est spécifique contrairement à l'immunité innée. Cette immunité adaptative est constituée de deux pans : celle de l'immunité à médiation cellulaire dans laquelle les lymphocytes T jouent un rôle central et celle de l'immunité à médiation humorale médiée par les lymphocytes B⁷. Les deux immunités agissent en synergie, après et avec la réponse innée, afin de détruire un organisme pathogène.

Afin d'exposer ici le principe sur lequel repose l'expression des immunoglobulines ainsi que les mécanismes impliqués dans la création de la diversité au niveau de ces dernières, je m'appuierais sur le site IMGT (ImmunoGeneTics). IMGT est considérée comme le **système d'information international** des immunoglobulines. Elle a été créée en 1989 par Marie-

Paule LEFRANC, présidente à l'institut universitaire de France de la partie immunogénétique et immuno-informatique.

II Les anticorps

II.1 Structure d'une immunoglobuline

En 1959, Rodney Robert Porter définit la structure d'un anticorps à la suite des travaux de Gérard Edelman. Les deux chercheurs obtiendront le prix Nobel en 1972⁸. Cette structure a été mise en évidence grâce à la cristallographie aux rayons X. Les anticorps appartiennent à la superfamille des immunoglobulines et sont des glycoprotéines. Une chaîne lourde (H pour heavy) est associée à une chaîne légère (L pour light) par un pont disulfure ainsi que des liaisons non covalentes telles que les liaisons hydrogènes, hydrophobes ou ioniques. Ceci forme alors un complexe H-L. Ce complexe H-L s'associant à un autre complexe H-L (strictement identique), par la présence de ponts disulfures au niveau de la région charnière de la région constante des IgG, IgA et IgD qui confère une flexibilité à la structure, forme une immunoglobuline (Ig) de base à quatre chaînes (**Figure 22**), notée $(H_2L_2)_n$ avec n variant de 1 à 5 selon la classe de l'Ig, dans les formes sécrétées. Chez l'homme, il existe 5 classes et 6 sous-classes d'immunoglobulines : IgM (μ), IgA ($\alpha 1$ et $\alpha 2$), IgE (ϵ), IgG ($\gamma 1$, $\gamma 2$, $\gamma 3$ ou $\gamma 4$) et IgD (δ). Pour les chaînes légères, il existe deux types : kappa (κ) et lambda (λ). Une chaîne légère est constituée d'un domaine variable (VL) et d'un domaine constant (CL). Une chaîne lourde est constituée quant à elle d'un domaine variable (VH) et de trois ou quatre domaines constants, notés CH1, CH2, CH3 (et CH4), selon l'isotype de la molécule. Le poids moléculaire d'une chaîne légère est de 25kDa, celui de la chaîne lourde de 50kDa. L'association H_2L_2 forme une structure de 150kDa.

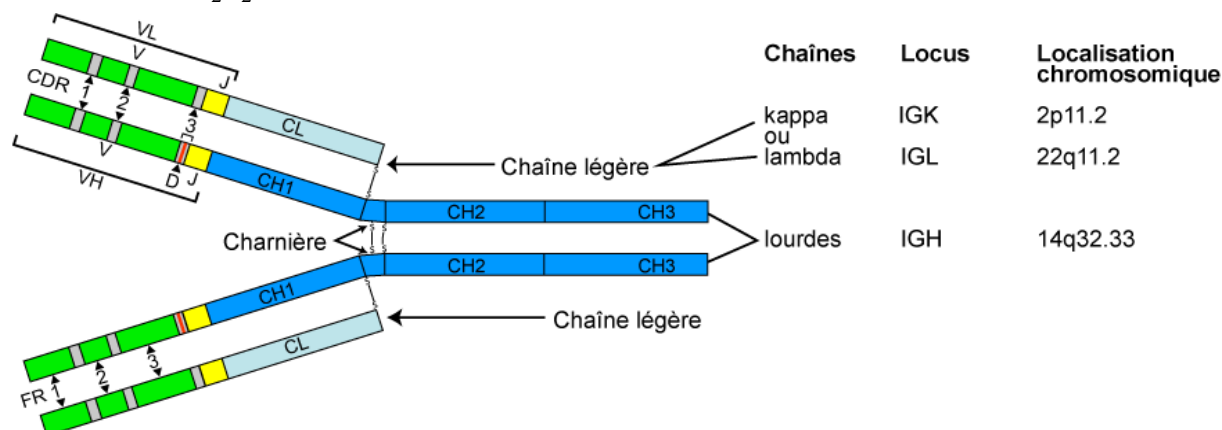


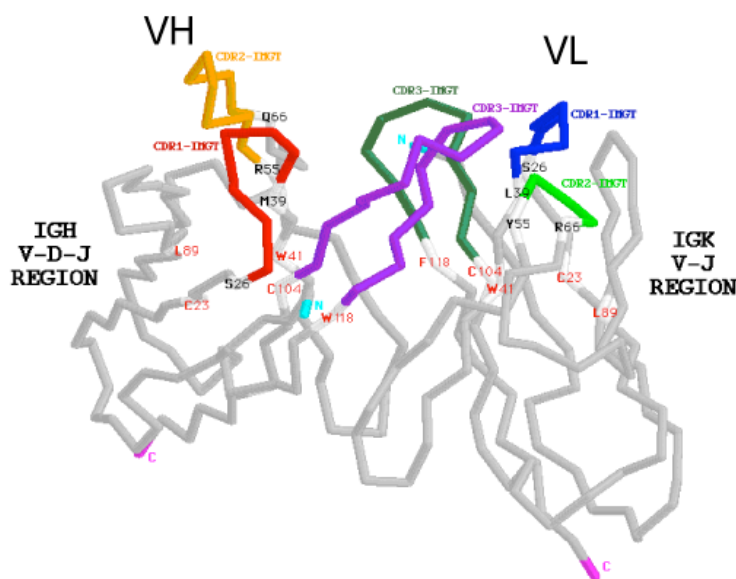
Figure 22 Représentation schématique de la structure d'une immunoglobuline, IgG1. Le domaine variable d'une chaîne légère (VL) est codé par le réarrangement de deux gènes, d'un V et d'un J spécifiques de la chaîne légère (kappa ou lambda). Le domaine variable d'une chaîne lourde (VH) est codé par le réarrangement de 3 gènes V, D et J. Dans chaque région variable, on trouve des régions hypervariables ou régions CDRs (complementarity Determining Region). Cette hypervariabilité est due à la variété combinatoire des différents gènes pour le CDR3 et à des hypermutations somatiques survenant lors de la forte prolifération des cellules B stimulées via leur BCR dans les OLS (organes lymphoïdes secondaires). Les 6 régions hypervariables (3 d'un domaine VH associées à 3 régions hypervariables d'un domaine VL) constituent le site de reconnaissance et de liaison à l'antigène. La région constante de la chaîne légère est codée par le gène IGKC (chaînes kappa) ou l'un des 7 gènes IGLC (chaînes lambda), et comprend un seul domaine constant. La région constante de la chaîne lourde est codée par des gènes IGHC et comprend 3 à 4 domaines constants. CH1, CH2 et CH3 représentent les domaines constants d'une chaîne lourde gamma. La région charnière située entre le domaine CH1 et CH2 des IgG est codée par 1 exon (cas des IgG1, IgG2 et IgG4 humaines) ou plusieurs exons, le plus souvent 4 (cas des IgG3) (Extrait du site IMGT, système d'information international des immunoglobulines)

II.2 Les domaines constants

Les domaines constants sont caractéristiques de l'espèce et de l'isotype. La fonction de la portion constante (Fc (CH₂ et CH₃)) réside d'une part dans l'activation du complément par son interaction avec le C1q. Cette lyse est nommée « CDC » (Complement dependent cytotoxicity) ou cytotoxicité dépendant du complément. La portion Fc est également impliquée dans l'élimination des complexes immuns anticorps/antigène par les cellules possédant des récepteurs au fragment Fc via la phagocytose (ADCP (Antibody dependent cellular phagocytosis)) mais également via un mécanisme de toxicité appelé « ADCC (Antibody dependent cell-mediated cytotoxicity) » due à une dégranulation après la reconnaissance des antigènes recouverts d'opsonines (= anticorps) ⁹.

II.3 Les domaines variables

Les domaines variables des Ig sont impliqués dans la liaison à l'antigène. L'association du domaine variable de l'une des chaînes lourdes (VH) au domaine variable de l'une des chaînes légères (VL) constitue l'un des sites d'interaction pour l'antigène. Il y a donc deux sites susceptibles de fixer deux antigènes sur une molécule d'anticorps. Les fragments variables des chaînes lourdes et des chaînes légères sont constitués de régions hypervariables au nombre de 3, nommées CDR (Complementary Determining Region) (**Figure 23**) et de régions conservées au nombre de 4, nommées FR (Framework). L'antigène se fixe préférentiellement sur les régions hypervariables et notamment au niveau du CDR3.



Copyright© IMG T the international ImMunoGeneTics information system®. All rights reserved.

Figure 23 Les domaines variables VH et VL des chaînes lourdes et légères d'une Ig ou d'un anticorps. VH :CDR1-IMGT est en rouge, CDR2-IMGT en orange et CDR3-IMGT en violet. VL :CDR1-IMGT est en bleu, CDR2-IMGT en vert et CDR3-IMGT en vert-bleu (Extrait du site <http://www.imgt.org>)

II.4 Production d'anticorps par les lymphocytes B

La réponse immunitaire se répartie en deux volets, l'immunité innée et l'immunité adaptative interagissant l'une avec l'autre. L'immunité innée, première ligne de réponse face à des agents pathogènes pénétrant dans l'organisme, déclenche immédiatement une inflammation; les

acteurs principaux sont les phagocytes (comprenant essentiellement les cellules dendritiques et les macrophages), les granulocytes (tels que les neutrophiles) et les cellules « natural killer » (NK). Le deuxième système est appelé immunité adaptative, il répond aux infections persistantes. Ce système coordonne une multiplicité d'acteurs, y compris ceux impliqués dans l'immunité innée. Parmi ces acteurs, les lymphocytes T et B sont des cellules clés de l'immunité adaptative déclenchant une coopération complexe. Ils régulent la réponse immunitaire globale. Les lymphocytes B, en particulier, ont un rôle unique en définissant la réponse humorale: les plasmocytes dérivés de cellules B sécrètent des immunoglobulines (Ig) dans le sang qui reconnaissent spécifiquement les agents pathogènes et déclenchent une réponse immunitaire coordonnée (fonction lytique des anticorps via le complément, phagocytose indirecte des pathogènes opsonisés par des anticorps par des phagocytes). La maturation des lymphocytes B à partir de la cellule souche hématopoïétique se produit dans la moelle osseuse, indépendamment de l'antigène (première partie de la **Figure 24**)¹⁰. L'évolution de la cellule souche en une cellule B immature implique le réarrangement des gènes des Ig codant pour le BCR. Les réarrangements des chaînes lourdes et chaînes légères sont illustrés en **Figure 25**. La première recombinaison se fait au niveau du locus IgH entre un gène (DH) et un gène (JH) donnant une cellule pro B. La jonction avec le gène (VH) conduit ensuite à une cellule dite « pré B ». Seuls les gènes VHDHJH en phase de lecture correcte et sans codon stop codent pour une région variable fonctionnelle (VpréBCR) et permettent la synthèse d'une chaîne lourde H(μ) intracytoplasmique. Une petite proportion de la chaîne μ est alors exprimée à la surface des cellules pré-B, en association avec une pseudo chaîne légère (ψ L). Cette chaîne correspond à l'association non covalente de deux polypeptides : un polypeptide de type V (nommé V pré B) et un polypeptide de type C (appelé protéine λ 5 chez la souris). Ce préBCR donne un signal de survie pour cette cellule B. Il va alors y avoir réarrangement au niveau de la chaîne légère et notamment au locus Ig κ entre un V κ et un J κ . Si aucun réarrangement n'est productif au niveau des deux allèles, le réarrangement se poursuit au niveau du locus IgL. Le remplacement de la pseudo-chaîne par une chaîne légère classique donne naissance à une IgM exprimée à la surface du lymphocyte B immature. Ces cellules produisent également un long transcrit préARNm couvrant les régions constantes des chaînes μ et delta (δ) qui par épissage alternatif va donner une cellule B naïve exprimant à la fois une IgM et une IgD à sa surface avant de sortir de la moelle osseuse. Elle est alors prise dans le flux sanguin pour se diriger vers les organes lymphoïdes secondaires (**Figure 24**). Là, le BCR peut entrer en contact avec son antigène spécifique, via une coopération avec les cellules T, qui déclenche l'activation des lymphocytes B, stade d'activation induisant une prolifération et une différenciation en plasmocytes (Cellules B effectrices) et en cellules B mémoires. Les plasmocytes sécrètent des immunoglobulines solubles d'isotype IgM, spécifiques à la réponse humorale primaire à court terme. Par commutation isotypique, il est possible que les cellules B matures expriment par la suite des immunoglobulines d'isotype G, E ou A. Les cellules B mémoires représentent alors une «base de données» conservée en mémoire des antigènes connus par l'organisme.

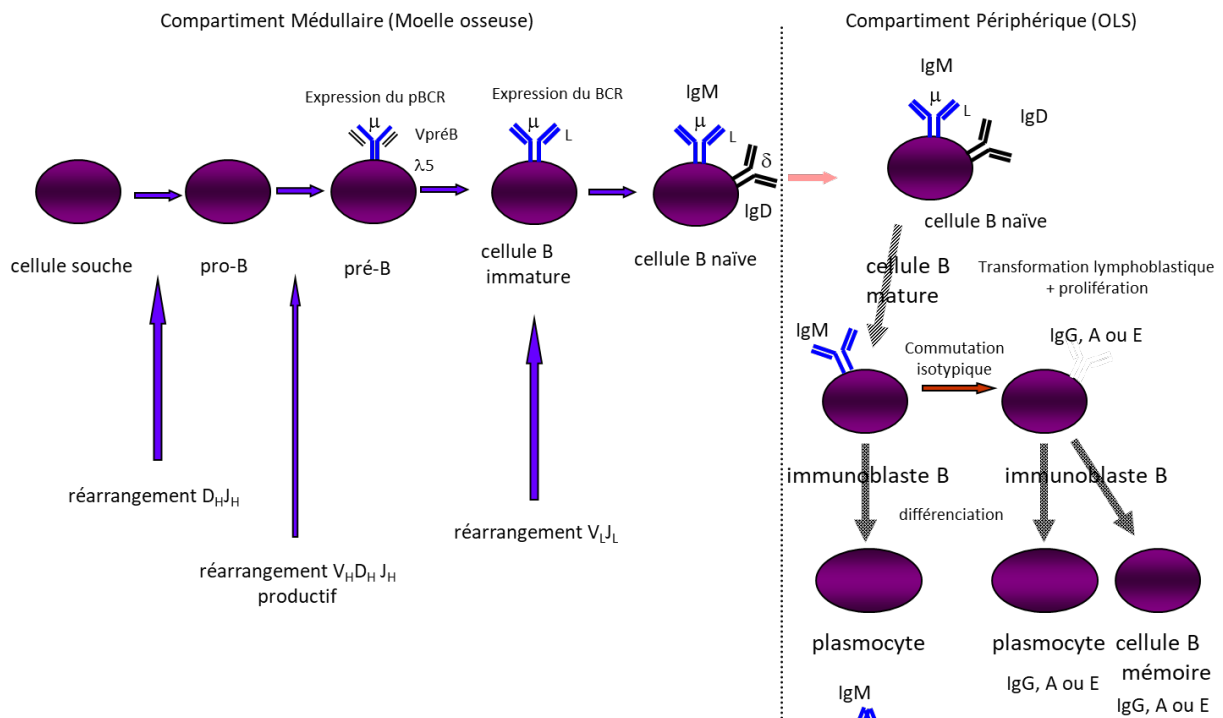


Figure 24 Production d'anticorps à partir des cellules B matures. La maturation du lymphocyte B à partir d'une cellule souche est indépendante de l'antigène et se produit dans la moelle osseuse. L'évolution de la cellule souche en une cellule B immature implique le réarrangement des gènes des Ig codant pour le BCR. La première recombinaison entre un gène (D_H) et un gène (J_H) donne une cellule pro B. La cellule proB devient une cellule préB après la jonction du gène (V_H) avec le DJ_H déjà réarrangé. Au niveau de cette cellule dite préB seuls les segments $V_HD_HJ_H$ correctement réarrangés et codant pour un VpréBCR permettent ensuite la synthèse d'une chaîne lourde H(μ) intracytoplasmique. Le préBCR correspond à la portion de la chaîne μ associée à une pseudo chaîne légère. Ce préBCR donne un signal de survie pour cette cellule B et entraîne le réarrangement au niveau de la chaîne légère et notamment au locus IgL entre un V_L et un J_L . La chaîne légère réarrangée vient remplacer la pseudo chaîne et permet l'expression d'une IgM à la surface du lymphocyte B immature. Un long transcrypt préARNm couvrant les régions constantes des chaînes μ et delta (δ) va par épissage alternatif donner une cellule B naïve exprimant à la fois une IgM et une IgD à sa surface avant de sortir de la moelle osseuse. Dans les OLS, la cellule B rentre en contact avec son antigène induisant une prolifération et une différenciation en plasmocytes et en cellules B mémoires. Les plasmocytes sécrètent des IgM mais qui par commutation isotypique pourront alors sécréter des IgA, E ou G. (Schématisation de l'ontogénie des lymphocytes B par Gisèle CLOFENT-SANCHEZ – directrice de l'équipe « ciblage de la plaque d'athérome »)

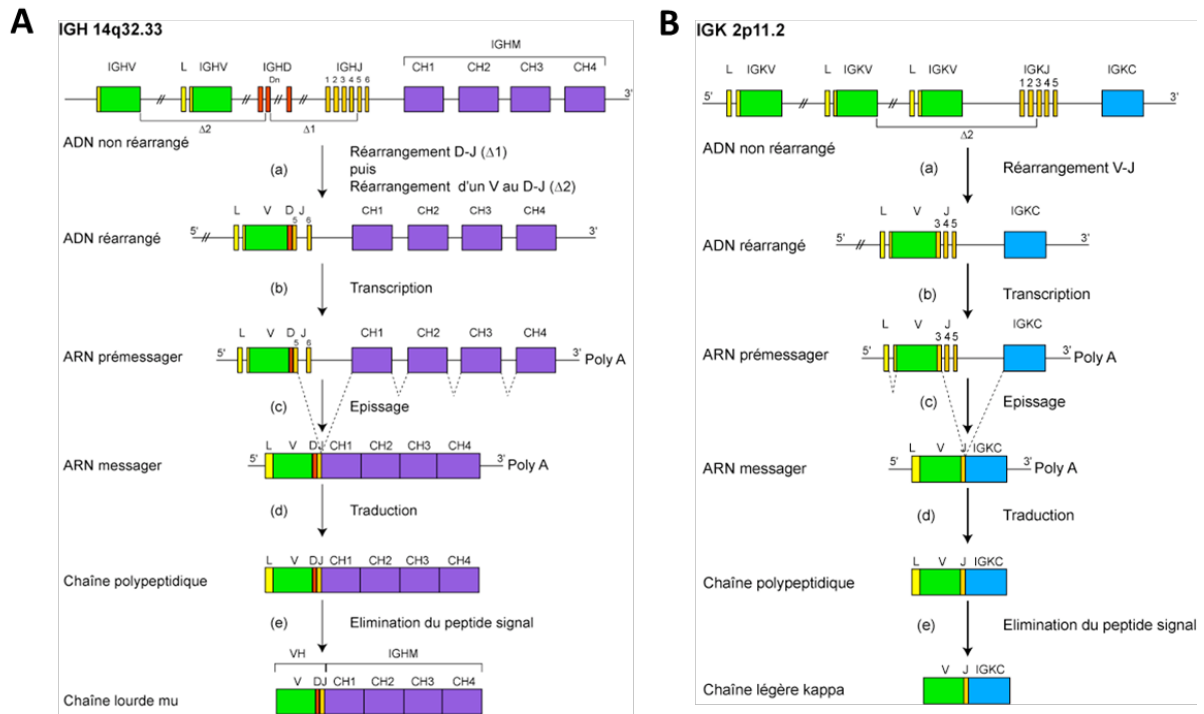


Figure 25 Synthèse d'une chaîne lourde mu (μ) et d'une chaîne légère kappa d'immunoglobuline. (A)(a1) Au niveau de l'ADN, lors d'une première étape l'un des gènes DH est joint à l'un des gènes JH avec délétion de l'ADN intermédiaire, pour créer un réarrangement D-J partiellement réarrangé. Dans une deuxième étape (a2), un des gènes VH est joint au D-J, avec délétion de l'ADN intermédiaire pour générer un gène VH-DHJH (IGHV-D-J) complètement réarrangé. (b) Le réarrangement IGHV-D-J est transcrit avec le gène IGHM en un pré-messager IGHV-D-J-M (ou IGHV-D-J-Cmu). (c) Les séquences d'ARN correspondant aux introns et aux gènes J non utilisés sont alors excisées par épissage et l'on obtient un ARN messager mature qui comprend les régions codantes épissées et les régions 5' et 3' non traduites. (d) L'ARN messager est ensuite traduit en une chaîne polypeptidique par les ribosomes. (e) Le peptide signal leader est éliminé par une peptidase après pénétration de la chaîne polypeptidique dans la cavité du réticulum endoplasmique. Une chaîne lourde mu (μ) est produite. (B)(a) Au niveau de l'ADN, l'un des gènes VL est joint à l'un des 5 gènes JL avec délétion de l'ADN intermédiaire, pour créer un réarrangement VL-JL (IGKV-)J. (b) Le gène IGKV-J est transcrit avec le gène IGKC en un ARN pré-messager IGKV-J-C. (c) Les séquences d'ARN correspondant aux introns et aux gènes IGKJ non utilisés sont alors excisées par épissage, et l'on obtient un ARN messager mature qui comprend les régions codantes épissées et les régions 5' et 3' non traduites. (d) L'ARN messager est ensuite traduit en une chaîne polypeptidique par les ribosomes. (e) Le peptide signal L est éliminé par une peptidase après pénétration de la chaîne polypeptidique dans la cavité du réticulum endoplasmique et une chaîne légère kappa mature est produite. (L: pour Leader) (Extrait et modifié à partir du site <http://www.imgt.org/IMGTEducation>)

III La diversité des anticorps

Les études de clonage cellulaire et de séquençage de l'ADN des chaînes lourdes et légères ont permis de comprendre la diversité et la complexité des anticorps, beaucoup plus importantes que celles prédites par Dryer et Bennett¹¹. En effet, en 1965, Dryer et Bennett ont suggéré que deux gènes séparés codaient pour la partie variable et pour la partie constante. Ces deux gènes doivent être réunis pour former un message continu qui sera transcrit pour donner l'une des deux chaînes, lourde ou légère. Cette théorie sera validée en 1976 par Tonegawa et Hozumi¹².

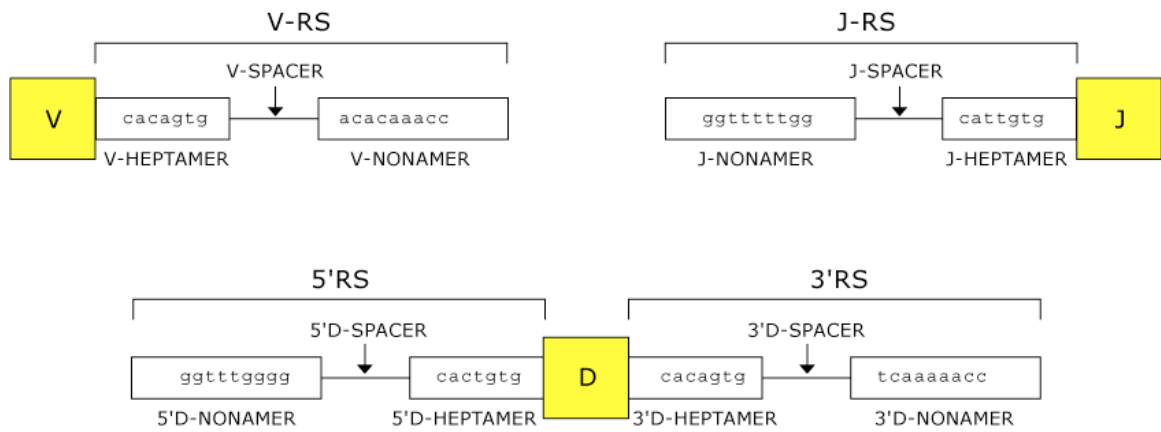
III.1 Les familles multigéniques codant pour les immunoglobulines et leur réarrangement

Le premier facteur de diversité repose sur l'existence de familles multigéniques de type V (Variable), D (Diversity) et J (Joining) qui sont les éléments constitutifs des domaines variables des chaînes lourdes et les chaînes légères κ et λ . Un gène V avec un gène J pris au hasard constitue la région variable d'une chaîne légère (VL), tandis qu'un domaine variable d'une chaîne lourde associe un V, un D et un J. Dans les cellules germinales, notamment au niveau de l'ADN, les familles multigéniques contiennent différents gènes séparés par des régions non codantes. Le tableau ci-dessous résume le nombre de gènes pour chaque famille V, D et J (**Tableau 3**). Le réarrangement des gènes de la lignée germinale dans les cellules somatiques s'effectue de manière aléatoire créant une augmentation de la diversité des anticorps. La diversité combinatoire créée par les réarrangements des gènes peut se calculer. En effet, les 38-46VH pouvant se combiner aléatoirement à l'un des 23 IGHD et des 6 IGJH, cela donne $38-46 \times 23 \times 6 \approx 6300$ combinaisons possibles ; les 31-35 IGHV $_{\kappa}$ pouvant se combiner aléatoirement à un des 5 IGLJ $_{\kappa}$, cela donne $31-35 \times 5 = 155-175$ combinaisons possibles au niveau du locus κ ; les 28-33V $_{\lambda}$ pouvant se combiner aléatoirement à un des 4-5J $_{\lambda}$, cela donne $28-33 \times 4-5 = 116-165$ combinaisons possibles au niveau du locus λ (Extrait du site <http://www.imgt.org/IMGTeducation>). De part et d'autre des exons V, D et J, il existe des séquences signal de recombinaison, dite RSS, qui orientent la recombinaison. Chaque séquence RS est constituée de deux motifs consensuels conservés de sept nucléotides (heptamère) et de neuf nucléotides (nonamère). Ces deux motifs sont séparés par des séquences peu conservées allant de 12 à 23 nucléotides appelées « spacer ». Ces « spacers » correspondent à un tour (RSS1) ou deux tours (RSS2) d'hélice, respectivement. Chaque segment V possède en son extrémité 3' une RSS, un segment J possède une RSS en 5' alors que les segments D en possèdent une en 5' mais également en 3' (**Figure 27**). La recombinaison se produit uniquement entre un RSS1 et un RSS2, garantissant la spécificité des réarrangements (pas de réarrangement entre deux segments V par exemple). La recombinaison, qui a pour but d'écarter les introns présents entre les gènes impliqués dans le réarrangement, est activée grâce à la présence d'enzymes RAG1 et RAG2 (Recombination-Activating Gene 1 et 2) (**Figure 26**).

Lors de la recombinaison, il se forme une épingle à cheveux qui, après coupure par les enzymes RAG 1 et RAG 2, peut être le lieu d'addition de P et N nucléotides. Les régions codées par les jonctions V-J et V-D-J pour les chaînes légères et pour les chaînes lourdes, respectivement, correspondent à la région CDR3.

Tableau 3 Nombre de gènes V, D et J pour les chaînes légères κ et λ et les chaînes lourdes

	Gènes V	Gènes D	Gènes J
Chaîne κ	31-35	-	5
Chaîne λ	29-33	-	4-5
Chaîne lourde	38-46	23	6



Copyright© IMGT the international ImMunoGeneTics information system®. All rights reserved.

Figure 27 Représentation graphique des séquences RS (3'V-RS, 5'J-RS, 5'D-RS et 3'D-RS). Les rectangles jaunes correspondent au gène V-3', 5'-J et au 5'-D-3' (Extrait du site <http://www.imgt.org>)

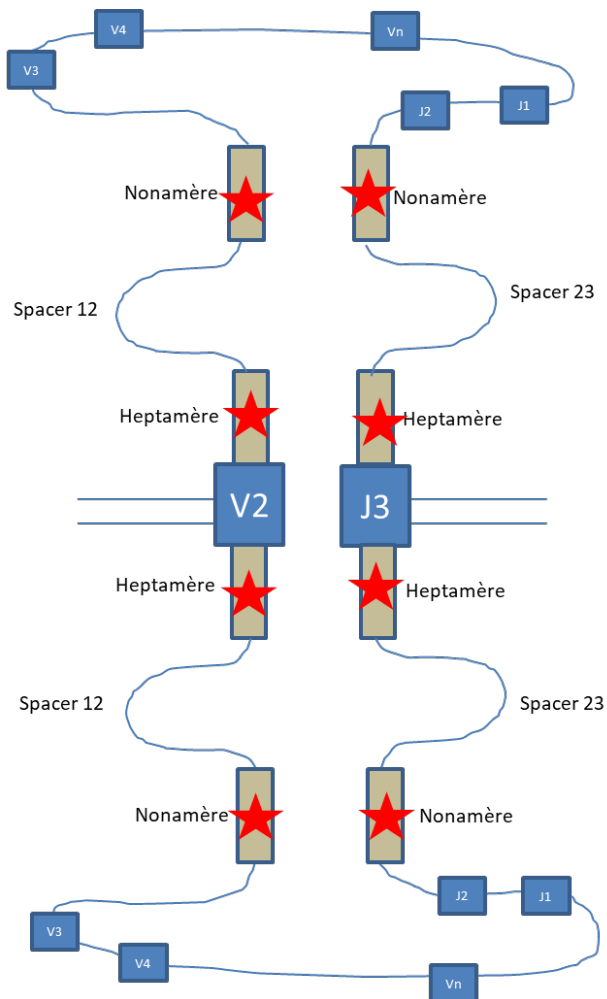


Figure 26 Représentation schématique de la formation de l'épingle à cheveux lors de la recombinaison V-J pour une chaîne légère. La recombinaison se fait entre un RSS1 (spacer 12) et un RSS2 (spacer 23). Les enzymes RAG1 et RAG2 vont reconnaître les associations des heptamères et nonamères (étoiles rouges) puis couper l'épingle à cheveux.

III.2 La flexibilité jonctionnelle

La diversité jonctionnelle permet d'augmenter encore la diversité créée par les mécanismes de recombinaison. Lors des processus de recombinaison V(D)J, les mécanismes de réparation de l'ADN créent une variabilité dans les zones de jonction entre les gènes associés et la position précise à laquelle les segments génétiques V(D)J se joignent peut légèrement varier. Ce phénomène induit un degré supplémentaire de diversité par délétion ou insertion de nucléotides dans les régions variables des immunoglobulines.

III.2.1 La P addition

Les insertions P concernent l'addition de quelques nucléotides complémentaires au niveau des joints codants. Cet ajout se fait par une polymérase après clivage asymétrique de l'épingle à cheveux par les endonucléases RAG1 et RAG2. (**Figure 28**).

III.2.2 La N addition

Les insertions N concernent l'addition jusqu'à 15 nucléotides au hasard par la terminal déoxynucléotidyl transférase (TdT) au niveau de la coupure de l'épingle à cheveux. Cette insertion est spécifique du stade précoce du développement du lymphocyte B, pendant lequel la TdT est exprimée et où a lieu la recombinaison V(D)J. La TdT ajoute ces nucléotides sans amorçage, avec une préférence pour des résidus G. Ces régions N sont ainsi riches en G-C (**Figure 28**).

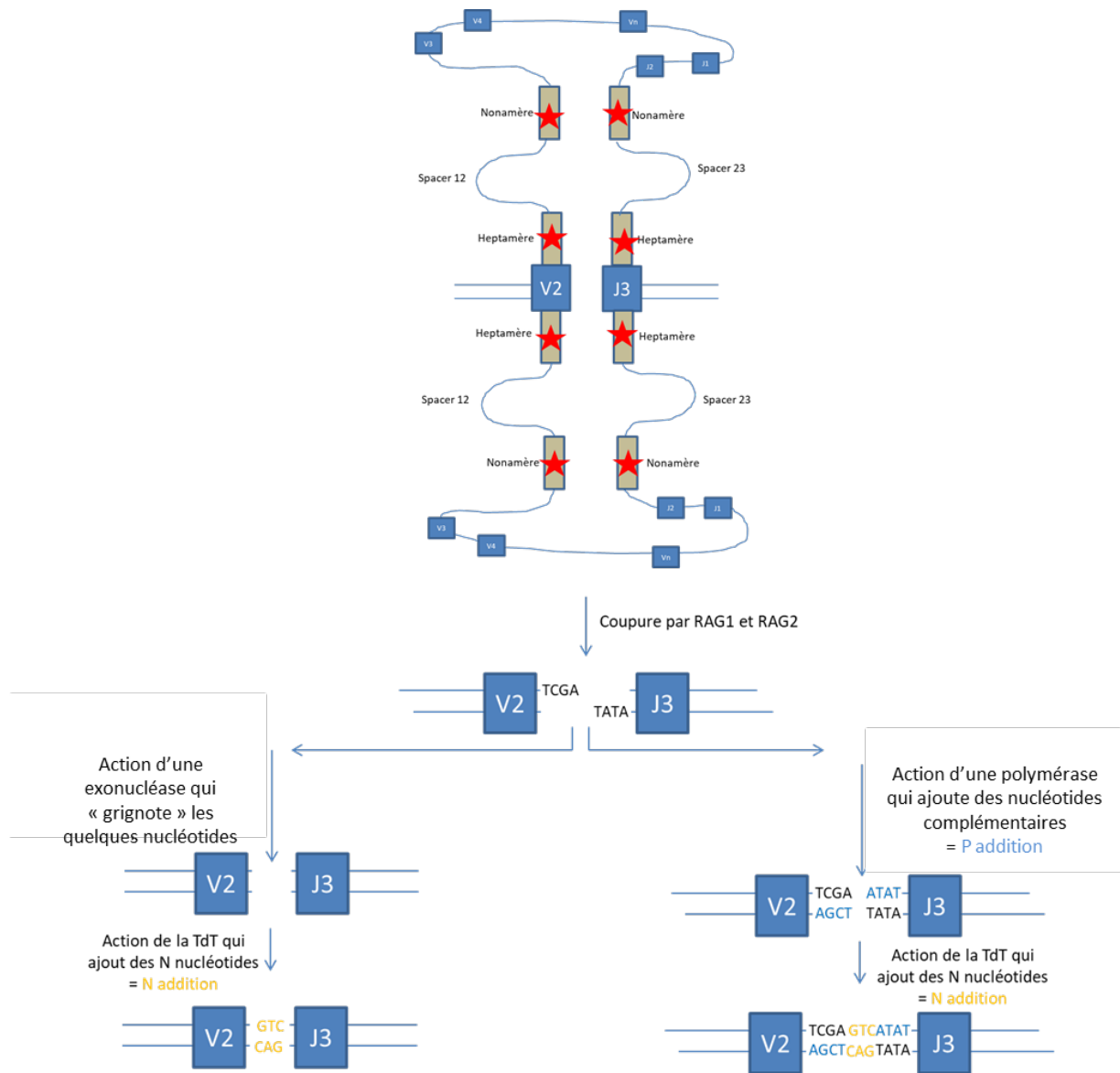


Figure 28 Schématisation de la P et N addition. Lors de la coupure par les enzymes RAG1 et RAG2 de l'épingle à cheveux, les extrémités sont souvent asymétriques. Dans ce cas-là, deux actions peuvent avoir lieu. La première consiste en une suppression des nucléotides supplémentaires par une exonucléase. La deuxième consiste en un ajout de nucléotides complémentaires par une polymérase. Il s'agit, dans ce cas, de la P addition. A la suite de l'une ou de l'autre des deux actions, il peut également y avoir un ajout aléatoire de nucléotides supplémentaires par la TdT au niveau de la jonction codante. Ce processus est appelé la N addition.

III.3 L'hypermutation somatique

Les modifications impliquées dans la diversité des anticorps expliquées ci-dessus s'opèrent lors des réarrangements V-J ou V-D-J dans la moelle osseuse lors de la maturation des lymphocytes B. Une diversité supplémentaire vient s'ajouter après réarrangements des gènes. En effet, certains nucléotides au niveau des régions variables sont remplacés par d'autres grâce à la cytidine déaminase (AID (Activated induced cytidine deaminase)), ce qui peut amener à la modification de l'affinité de l'immunoglobuline (Extrait du site <http://www.imgt.org/IMGTEducation>). Les hypermutations somatiques ont lieu dans les organes lymphoïdes secondaires lors de l'activation des clones B matures par l'antigène en particulier lors de la forte prolifération des centroblastes dans la zone sombre des follicules. Elles se font donc chez les cellules B matures essentiellement au niveau des CDR1, 2 et 3 du domaine VH mais également dans les CDR1, 2 et 3 du domaine VL. Suite à une exposition à

l'antigène, les cellules B présentant les récepteurs de plus forte affinité vont survivre préférentiellement. En résulte une augmentation de l'affinité appelée maturation de l'affinité.

III.4 Association aléatoire des chaînes lourdes et légères

Comme vu précédemment, chez l'homme il existe environ 6300 combinaisons possibles au niveau des chaînes lourdes VH et jusqu'à 340 combinaisons possibles au niveau des chaînes légères VL. Par hypothèse, n'importe lequel des gènes des chaînes lourdes et n'importe lequel des gènes des chaînes légères peuvent être présents simultanément dans la même cellule ; le nombre de combinaisons possibles est donc de 2 142 000.

III.5 Conclusions

Il existe donc trois niveaux de diversité : (1) la combinatoire aléatoire des segments V-J et V-D-J, (2) la flexibilité jonctionnelle et (3) l'hypermutation somatique. Ces trois niveaux de diversité additionnés à la combinaison aléatoire des chaînes lourdes avec les chaînes légères créent une diversité des anticorps chez l'homme supérieure à 10^{12} , permettant ainsi au répertoire de reconnaître un très grand nombre d'antigènes différents (Extrait du site <http://www.imgt.org/IMGTeducation>).

IV Les différents formats d'anticorps

L'immunoglobuline d'isotype G est l'immunoglobuline la plus retrouvée dans l'organisme (concentration comprise entre 0,5 et 16 mg/mL). Elle contient plusieurs domaines protéiques susceptibles d'être séparés par l'utilisation d'enzymes telle que la papaïne ou la pepsine mais également grâce au progrès du clonage et la manipulation de l'ADN (**Figure 29**). Ainsi, grâce à l'ingénierie des anticorps, différents formats peuvent être directement synthétisés¹³. Certains de ces formats posant des problèmes de stabilité tels que les Fv (association d'un VH et d'un VL), d'autres formats ont été conçus. Ainsi, l'un des formats les plus connus est le scFv qui correspond à l'association d'un VH relié à un VL grâce à une courte séquence peptidique nommée « linker ». Le scFv possède plusieurs caractéristiques qui en font un format de choix notamment sa taille inférieure à 30kDa et sa capacité à pénétrer dans les tissus¹⁴. De par leur format monovalent, les scFv montrent assez fréquemment des constantes de dissociation rapides et des temps de rétention faibles sur l'antigène ciblé. De ce fait, les chercheurs ont développé le format scFv-Fc correspondant à l'association de deux scFv grâce à l'ajout d'un fragment constant Fc, souvent muté pour éviter la reconnaissance par les RFcs. La bivalence ainsi créée augmente alors l'avidité du scFv pour l'antigène¹⁵. Un autre format d'anticorps très intéressant en vue d'un objectif thérapeutique car ayant une séquence en acides aminés très proche de celle retrouvée chez les VH de l'homme est le VHH, nommé aussi « nanobody ». Ce format est le plus petit format dérivé des anticorps simple chaîne découverts chez les camélidés comme le lama ou le chameau. Non associé à un VL, la reconnaissance de l'antigène s'effectue uniquement par un seul domaine variable. L'ingénierie des anticorps ouvre aujourd'hui la voie à l'obtention de formats extrêmement intéressants pour l'imagerie et la thérapie dans le cadre de diverses pathologies comme les maladies auto-immunes, les cancers.

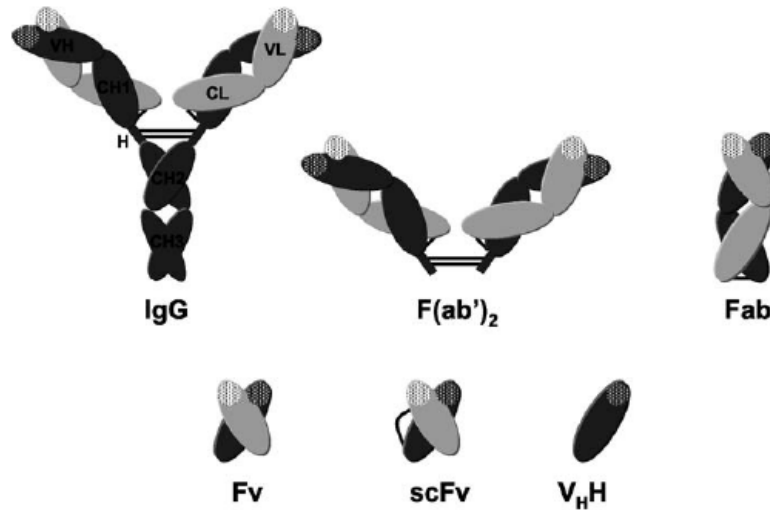


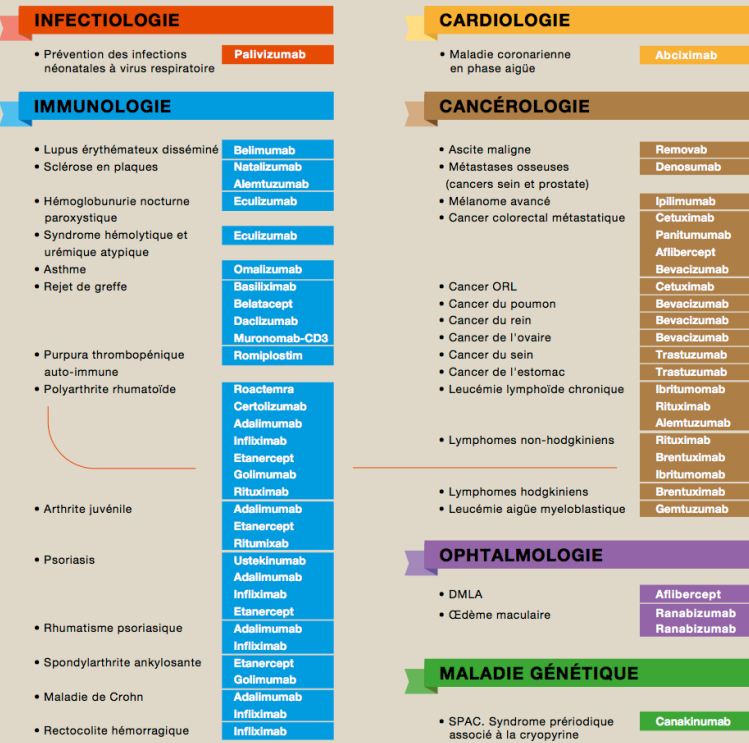
Figure 29 Différents formats d'anticorps. Les fragments Fab et F(ab')₂ sont généralement obtenus respectivement par digestion à la papaïne et la pepsine d'une IgG entière. Le Fv correspond à l'association non covalente d'un domaine variable VH à un domaine variable VL. Le scFv est une molécule monocaténaire où le VH est relié au VL par un peptide de liaison dit « linker ». Le V_HH est le plus petit domaine variable connu capable de lier un antigène avec une bonne affinité. Les cercles contenus dans les domaines VH, V_HH et VL schématisent les régions hypervariables appelées CDRs. (Extrait de la revue de Baty et Chames. paru en 2006)

V Les différents systèmes de production d'anticorps

Les anticorps monoclonaux (Acm) constituent aujourd'hui la classe thérapeutique la plus en développement dans le secteur pharmaceutique avec un taux de croissance annuel de 14% dans un marché mondial estimé à 42 milliards de dollars en 2010, ce qui représente 40% du marché du biomédicament. Selon la Leem (Les Entreprises du Médicament) consulté le 12/10/17, plus d'une trentaine d'anticorps monoclonaux sont commercialisés et près de 350 sont en développement clinique dont 45% dans les domaines de la cancérologie, 6% en pneumologie, 5% en gastroentérologie et 5% en dermatologie (**Figure 30**). En 2008, le besoin en Acm dépassaient déjà les 2 tonnes ce qui impose d'avoir des techniques de production efficaces et rentables. Plusieurs systèmes sont actuellement disponibles et en cours de développement même si le système de production le plus utilisé est le système mammalien notamment grâce aux cellules CHO (chinese hamster ovarian). Le tableau suivant résume les différents systèmes existants qui seront détaillés ci-après (**Figure 31**).

Bientôt des anticorps monoclonaux (AcMo) pour toutes les maladies ?

ANTICORPS MONOCLONAUX THÉRAPEUTIQUES : RÉPARTITION PAR MALADIES



35

C'est le nombre d'anticorps monoclonaux (AcMo) à la disposition des patients aujourd'hui¹

355

C'est le nombre d'anticorps monoclonaux en développement¹

- dont
- 45 % en cancérologie
 - 6 % en pneumologie
 - 5 % en gastroentérologie
 - 5 % en dermatologie

1

C'est le nombre d'anticorps monoclonaux anticancéreux approuvés chaque année par les agences de santé¹

Source : Hervé Watier. Présentation «Médicaments et biosimilaires : des médicaments d'avenir pour l'industrie pharmaceutique», LabEx «MABImprove», 6 août 2013

© LEEM - LE 100 QUESTIONS - 2015

Figure 30 Tableau des anticorps présents actuellement sur le marché (Extrait du site <http://www.leem.org/article/bientot-des-anticorps-mono-clonaux-acmo-pour-toutes-les-maladies>)

	Temps de génération du clone producteur ¹	Capacité de production ²	Glycosylation ³	Activité ADCC ⁴	Sécurité biologique ⁵
Bactéries	++++	+	-	-	++++
Levures	++++	++	+	-	++++
Champignons filamenteux	++	++	+	-	++++
Cellules d'insectes	++	+	++	+	++++
Plantes transgéniques	+	++++	++	++	++++
Animaux transgéniques	+	++++	+++	++++	+++
Cellules aviaires	++	++	++++	++++	++++
Cellules de mammifères	++	+++	++++	++	+++

Figure 31 Avantages comparatifs des systèmes de production d'anticorps monoclonaux (Extrait du site http://ipubli-inserm.inist.fr/bitstream/handle/10608/6772/MS_2009_12_1163.html#R1)

V.1 La glycosylation des anticorps

Idéalement, le système de production doit permettre de produire des anticorps présentant les mêmes caractéristiques pharmacologiques et pharmacocinétiques que les anticorps naturels. Ces propriétés biologiques étant contrôlées, en grande partie, par le profil de glycosylation

(**Figure 32**), les systèmes mammifères ont été pendant longtemps préférés à d'autres systèmes alternatifs du fait de la présence de la machinerie de glycosylation permettant des modifications post-traductionnelles proches de celles retrouvées chez l'homme. La glycosylation, qui se produit lors du passage dans le réticulum endoplasmique, est essentielle car elle conditionne le temps de demi-vie dans la circulation sanguine ainsi que l'interaction aux cellules effectrices. Ainsi, une modification des N-glycanes peut d'une part affecter l'activité biologique ainsi que la stabilité et la solubilité des anticorps mais également engendrer une immunogénicité des anticorps (notamment par la présence de $\alpha 1,3$ fucose). L'analyse des anticorps et notamment de la relation structure/fonction a permis de mieux comprendre le rôle des N-glycanes dans l'interaction avec les cellules. L'absence de fucose sur les structures N-glycaniques augmente l'activité cytotoxique des IgG1^{16,17}.

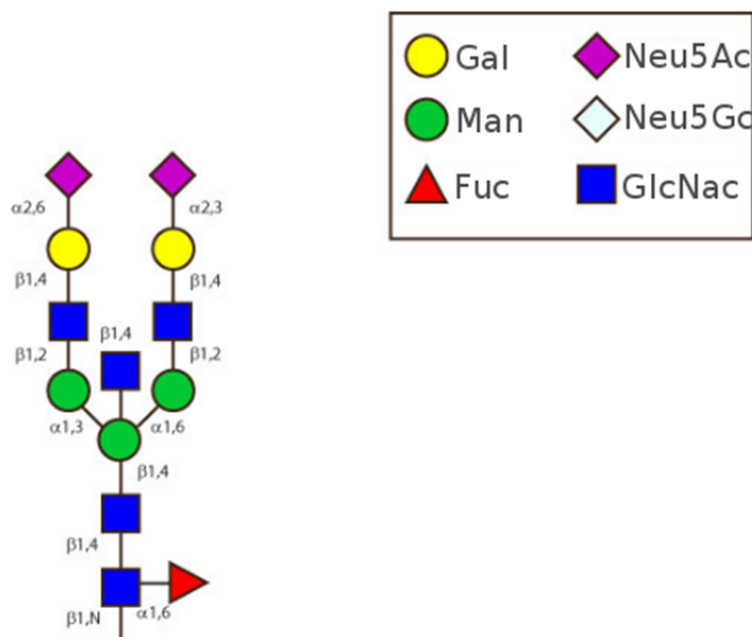


Figure 32 Profil de glycosylation des anticorps naturels. Gal: galactose, Man: mannose, Fuc: fucose, GlcNAc: N-acetylglucosamine, Neu5Ac: N-acetylneuraminic acid and Neu5Gc: N-glycolylneuraminic acid. (<http://www.imgt.org/IMGTeducation/IMGTlexique/G/Glycosylation.html>)

V.2 Le système mammifère

La lignée cellulaire CHO est utilisée depuis deux décennies pour la production d'anticorps. Ce système permet de synthétiser des molécules présentant des caractéristiques chimiques et biologiques proches des molécules synthétisées naturellement. De plus, des progrès technologiques au niveau des vecteurs d'expression ainsi que la maîtrise des procédés industriels a permis de passer de quelques milligrammes d'anticorps par litre de milieu de culture à des grammes par litre. L'un des points importants à prendre en compte également est notamment la glycosylation des anticorps (comme discuté précédemment) avec des modifications post-traductionnelles qui assurent en grande partie la fonction de l'anticorps. Les cellules humaines telles que les cellules de rétine Per.C6, les cellules embryonnaires (HEK (Human embryonic kidney)), les cellules cancéreuses du col de l'utérus ou encore les lymphocytes T (Jurkat) ont fait leur apparition et présentent l'énorme avantage de la glycosylation humaine. Cette glycosylation permet ainsi l'interaction avec les cellules effectrices et est essentielle au temps de demi-vie des anticorps dans la circulation sanguine. Cependant un inconvénient majeur du système de production mammalien étant son coût très onéreux, des systèmes alternatifs ont été développés.

V.3 Le système bactérien

Grâce à de meilleures connaissances dans le domaine de la biologie moléculaire, le système de production dans *Escherichia Coli* est devenu le système de production permettant d'obtenir à grande échelle des quantités importantes de protéines recombinantes en raison notamment de la facilité et de la rapidité de génération de clones producteurs et de coûts de production avantageux. Cependant, la bactérie étant incapable de produire des anticorps correctement repliés et glycosylés et donc dépourvus d'activités effectrices, ce système est utilisé principalement pour la production de fragments d'anticorps comme les Fv ou scFv (single chain fragment variable) mais également pour la production d'anticorps mono-chaînes V_HH comme le fait la société Ablynx fondée en 2001. De ce fait, actuellement aucun anticorps monoclonal entier produit dans le système bactérien n'est présent sur le marché du médicament.

V.4 Le système levure et champignon filamenteux

Tout comme le système bactérien, les levures et champignons filamenteux offrent une forte production avec des coûts de production faible, une maîtrise de la technologie ainsi qu'une présence sur le marché des produits approuvés par les autorités de santé. Le système levure présente un avantage supplémentaire par rapport à *E.Coli* qui est d'avoir une machinerie de glycosylation permettant la N-glycosylation ainsi que l'addition de ponts disulfures sur des protéines complexes (exemple des anticorps entiers). La production d'anticorps se fait notamment dans *Pichia pastoris* et *Aspergillus niger*, ces deux souches ayant la capacité de réaliser une N-glycosylation complexe mais caractérisée par la présence de structures riches en mannoses. Les recherches actuelles tentent de réduire ce problème de glycosylation par humanisation du système de glycosylation en éliminant les gènes contrôlant l'hypermannosylation et en introduisant des gènes catalysant la synthèse et l'addition de glycanes humains. Cette nouvelle perspective d'humanisation du système de glycosylation étant prometteuse et pouvant apporter de nombreuses solutions, la société Merck a racheté pour 400 millions de dollars la société américaine Glycofi.

V.5 Le système des cellules d'insectes

De même que les levures, les cellules d'insectes (Sf9 et Sf21, les plus utilisées en laboratoire) infectées par un vecteur issu du baculovirus sont capables de produire des protéines complexes glycosylées. Cependant, il a été observé que les structures glycaniques portées par les protéines recombinantes sont en général tronquées. Parmi les glycanes tronqués ou les glycanes « non humains » se trouve le α 1,3-fucose qui augmente l'immunogénicité de la molécule. La question suivante s'est donc posée : comment humaniser le système de glycosylation ? Parmi les équipes qui s'y sont intéressées, l'équipe de Martine CERUTTI (UPS 3044) travaillant sur le système baculovirus a proposé une approche d'humanisation. Ainsi, il s'agit de modifier le patrimoine enzymatique de la cellule par l'addition de glycosyltransférases ou encore d'inactiver des gènes responsables de la biosynthèse de glycanes indésirables. Il a été démontré que les cellules d'insectes sont capables de faire de la N-glycosylation (l'oligoside N-acétylglucosamine (GlcNAc) se lie alors au résidu asparagine libre) et de la O-glycosylation (l'oligoside N-acétyl-galactosamine est, dans ce cas, ajouté sur les résidus -OH des sérines et thréonines). Les phases précoces de la voie de biosynthèse des N-glycanes sont identiques à celle des mammifères avec la synthèse de GlcNAcMan₅(Fuc)GlcNAc₂¹⁸. Néanmoins, il n'y a pas de sucres complexes (pas de galactose, et pas d'acide sialique). Des approches ont donc été mises au point dans le laboratoire de Martine CERUTTI en apportant des glycosyltransférases manquantes ou déficientes à la cellule d'insecte pour qu'elle puisse synthétiser des glycoprotéines galactosylées et/ou sialylées. Des glycosyltransférases impliquées dans la biosynthèse des glycanes complexes ont été clonées comme la Fut8 (α 1,6 fucosyltransférase)¹⁹ et la GNT-I

(glucose N-acetyltransferase). La complémentation des cellules d'insectes a été faite grâce à une modification du baculovirus. Un autre problème majeur de ce système est l'activité lytique post infection virale ce qui nécessite d'avoir à disposition un grand nombre de virus ainsi qu'un grand nombre de campagnes d'infection et de production. De manière à contourner le dernier problème de l'activité protéasique, la société Chesapeake PERL a démontré comment infecter directement les larves d'insectes par voie orale et purifier les protéines recombinantes directement à partir de ces dernières. Ainsi, cette société a réussi à produire des fragments Fab dans les cellules d'insectes *Trichoplusia ni*²⁰.

V.6 Les plantes transgéniques

Les plantes étant facilement modifiables génétiquement avec une conservation de clones producteurs peu contraignante, des anticorps monoclonaux ont été produits chez cet hôte. Ainsi en 1989, un anticorps monoclonal a été produit dans *Nicotiana tabacum* après transfection par le vecteur *agrobacterium tumefaciens*²¹. Comme chez les cellules d'insectes, la présence de résidus glycanes toxiques et immunogènes poussent à l'humanisation du système de glycosylation également chez les plantes transgéniques. Cependant, ce système de production est extrêmement remis en question à l'heure actuelle à cause d'un problème d'opposition publique suscitée par les risques de dissémination potentielle des plantes transgéniques cultivées en plein champ. Pour pallier ce problème, les cellules de plantes cultivées en bioréacteur en milieu clos font l'objet d'un intérêt croissant. La société américaine Biorex Therapeutics a ainsi développé un système d'expression basé sur la culture des lentilles d'eau *Lemna minor*. La société a également réussi à humaniser la machinerie de glycosylation de la plante *Lemna minor* grâce à des ARN interférents bloquant la synthèse des enzymes β 1.2-xylosyltransférase et α 1.3-fucosyltransférase endogènes. Il en est de même pour la société Greenovation Biotech qui a développé une plateforme avec la mousse *Physcomitrella patens* pour une production d'anticorps monoclonaux montrant des profils de glycosylation identiques à l'homme.

V.7 Les animaux transgéniques

La production d'anticorps monoclonaux peut s'effectuer dans le lait d'animaux transgéniques. C'est ce système qu'a développé la société GTC biotherapeutics, une spin-off de Genzyme. D'un point de vue historique la société GTC a été créée pour travailler sur la transgénèse et la possibilité de produire des anticorps dans les animaux transgéniques et en particulier dans le lait. Par la suite, cette société a été rachetée par la filiale LFB, LFB Biotechnologies, qui l'a alors renommée rEVO. La société a été ensuite divisée en deux : rEVO valorise ce qui est déjà existant alors que le LFB USA détient la partie innovation dédiée à la transgénèse^{22,23}.

VI Les anticorps en thérapie

Les anticorps sous leurs différents formats sont aujourd'hui en pleine expansion dans le secteur pharmaceutique. De par leurs propriétés notamment en termes d'affinité, de spécificité et de stabilité, ces molécules sont des agents thérapeutiques de choix. En vue de réduire leur immunogénicité et d'accroître leur efficacité, les anticorps thérapeutiques ont évolué d'un format murin jusqu'à un format totalement humain (**Figure 33**).

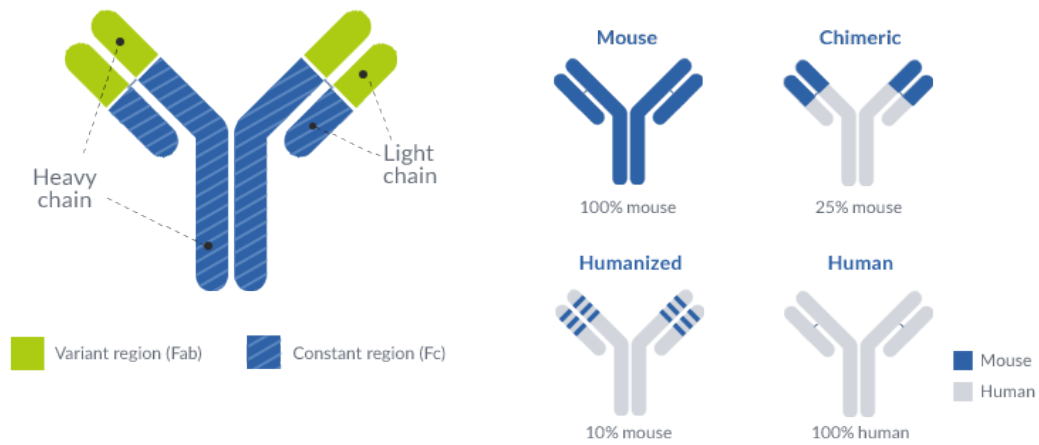


Figure 33 Structure de l'anticorps humain et représentation des distinctions entre chaque type d'anticorps: Souris, Chimérique, Humanisé et Humain (Extrait du site <https://www.proteogenix.science/antibody-production/phage-display-services/monoclonal-antibody-humanization/>)

VI.1 Les anticorps murins (-momab)

En 1975, Milstein et Kohler font une découverte qui leur permettra d'obtenir le prix Nobel de médecine et de physiologie en 1984. En effet, ces deux chercheurs ont développé une méthode élégante pour la production continue *in vitro* d'anticorps monoclonaux par des lymphocytes B immortalisés. Ce développement s'est basé sur la connaissance de ces deux chercheurs des cellules capables de produire en continu des anticorps identiques dans le cas de pathologies prolifératives (cas du myélome en particulier)²⁴. La technique des hybridomes est illustrée ci-dessous (**Figure 34**). Brièvement, une souris est immunisée avec l'antigène d'intérêt. Après immunisation, les lymphocytes B sont récupérés puis fusionnés en présence de PEG à des cellules myéloïdes. La fusion de ces deux cellules génère des hybridomes qui, après criblage pour obtenir les cellules B exprimant l'anticorps dirigé contre la protéine d'intérêt, sont capables de produire en continu ces anticorps dit « monoclonaux ».

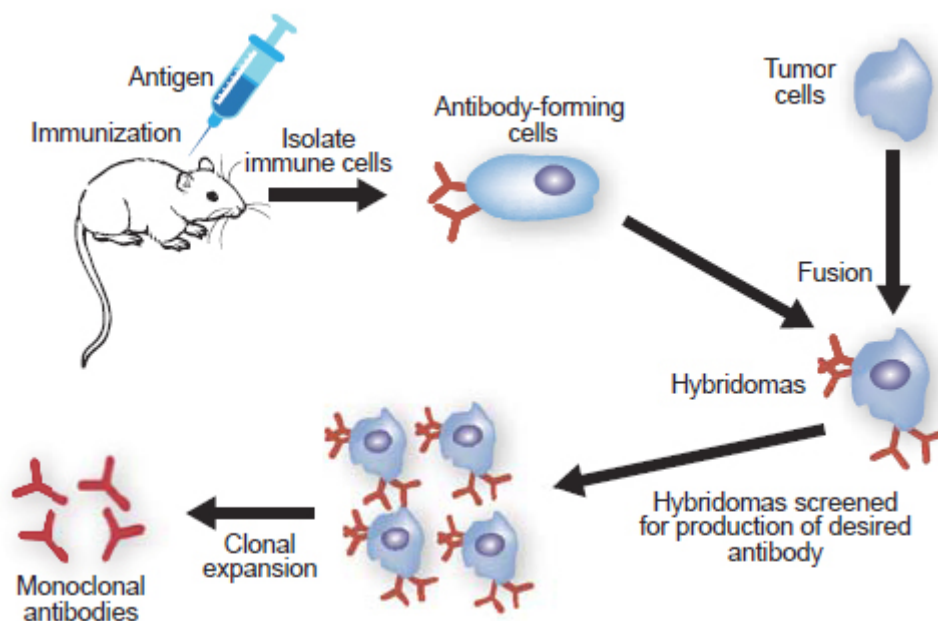


Figure 34 La technique des hybridomes implique la formation de lignées cellulaires dites « hybrides » (appelées hybridomes) en fusionnant une cellule B produisant des anticorps avec une cellule de myélome (cancer des cellules B) qui est choisie pour son incapacité à produire sa propre protéine de myélome (Extrait de <https://www.genscript.com/hit-generation-screening.html>)

Ainsi, des anticorps murins monoclonaux ont pu être produits par la technique des « hybridomes » comme le **Muromomab** (anti-CD3) pour être utilisés à des fins de diagnostic et de recherche. Cependant l'un des problèmes majeurs étant leur immunogénicité par induction d'une réponse HAMA (human anti mouse antibody) lorsqu'ils sont injectés chez l'homme²⁵, les recherches se sont tournées vers une modification des anticorps murins avec dans un premier temps la production d'anticorps chimériques.

VI.2 Les anticorps chimériques (-ximab)

Un anticorps chimérique correspond au clonage en fusion des régions variables murines (« donneur ») avec les régions constantes humaines (« accepteur »)²⁶. Ainsi, il est possible de conserver la spécificité et l'affinité du « donneur » tout en diminuant l'immunogénicité. En thérapie, la chimérisation constitue le premier pas vers l'humanisation de l'anticorps. Dans le cadre de l'athérosclérose, des anticorps chimériques ont été générés. Le premier anticorps approuvé par la FDA en 1994 est le **ReoPro**, anticorps chimérique dirigé contre l'intégrine $\alpha 2b\beta 3$, intégrine exprimée à la surface des plaquettes impliquées dans les événements thrombotiques. Cependant, au cours d'études sur des cohortes de patients, il a été montré que certains patients ayant reçu cet anticorps développaient des thrombocytopenies^{27,28,29}. Par la suite, d'autres anticorps chimériques ont été développés notamment un anticorps chimérique chP3R99 se liant aux GAG (GlycosAminoGlycanes) interférant donc avec la fixation des LDL aux GAG³⁰. L'hypothèse était qu'une vaccination par cet anticorps pourrait générer une captation des LDL par une cascade d'anticorps anti-idiotypiques dirigés contre les GAG. L'injection de cet anticorps chimérique chez des rats Sprague-Dawley diminuait la rétention des LDL au niveau de l'endothélium et leur oxydation. Dans la même étude, il a été montré que cet anticorps injecté chez le lapin New-Zealand prévenait le développement de la plaque d'athérome après induction au Lipofundin (émulsion de lipides produite par B.Braun Medical)³¹. D'autres anticorps chimériques sont à l'heure actuelle sur le marché des anticorps notamment le **Rituximab** (anti-CD20) ou encore le **Cetiximab** (anti-EGFR).

VI.3 Les anticorps humanisés (-zumab)

Un anticorps humanisé correspond à l'insertion des CDRs murins d'un anticorps « donneur » dans une charpente humaine d'un anticorps « accepteur »³². Les CDRs étant les sites d'interaction avec l'antigène, il est ainsi possible de conserver l'affinité de l'anticorps tout en diminuant de manière significative l'immunogénicité. De nos jours, la plupart des anticorps utilisés en thérapie et acceptés sur le marché sont des anticorps humanisés. Citons par exemple le **Trastuzumab** (anti-HER2 (Human Epidermal Growth Factor Receptor-2)) prescrit dans le cancer du sein. Dans le cadre de l'athérosclérose, le dernier anticorps humanisé développé est le **Bococizumab** (anticorps anti-PCSK9) par Pfizer. Cependant, Pfizer a annoncé que son développement a été interrompu. La raison est la suivante : je cite « les données cliniques montrent qu'il est peu probable que ce nouveau médicament apporte une valeur ajoutée aux patients, aux médecins ou aux actionnaires ». (<https://francais.medscape.com/voirarticle/3602784>).

VI.4 Les anticorps recombinants humains (-mumab ou -umab)

De par les problèmes d'immunogénicité et de spécificité, les recherches se tournent aujourd'hui vers le développement d'anticorps entièrement humains. Pour cela, deux techniques subsistent. La première méthode de production d'anticorps monoclonaux est la technique des souris transgéniques (**Figure 35**)³³. Après immunisation de souris nommées « xenomouse » dont les loci codant pour la production d'anticorps murins ont été inactivés et remplacés par des gènes humains portés par des YAC (Yeast Artificial Chromosome), les gènes humains se réarrangent et subissent des mutations somatiques aboutissant à la

production d'anticorps humains de forte affinité. La deuxième technique est la méthode du phage display qui sera discutée dans la section suivante. Parmi les anticorps humains obtenus par phage display et utilisés à l'heure actuelle en médecine, citons le **Necetimumab** dirigé contre l'EGFR (Epidermal Growth Factor Receptor) et prescrit dans le cadre du carcinome hépatique ou encore **Belimumab**, qui est un anti-BLyS (B lymphocyte stimulator) prescrit dans le cadre de la sclérose en plaque.

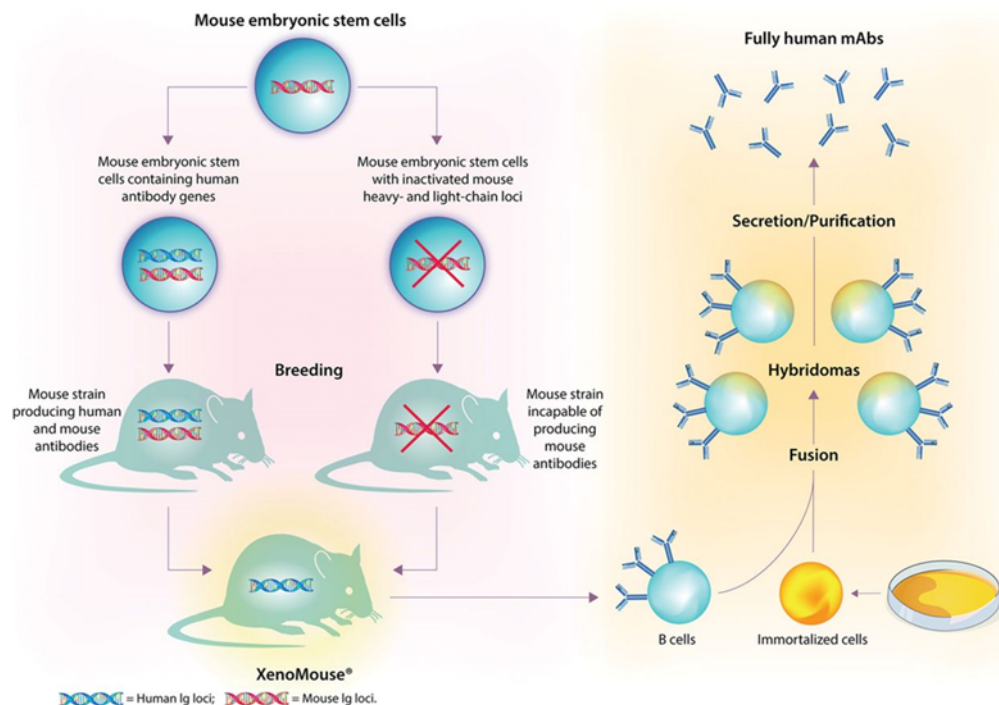


Figure 35 Souris transgéniques produisant des immunoglobulines humaines, exemple des Xenomouse®. L'obtention de xénomouse se fait par remplacement des gènes souris codant pour les Ig par leurs équivalents humains. Les plasmocytes de la Xenomouse sont ensuite immortalisés par la technique des hybridomes pour produire des anticorps monoclonaux entièrement humains (mAbs) (Extrait de l'article de Folz et al. paru en 2013 dans *Circulation*)

VI.5 Applications thérapeutiques des anticorps

L'utilisation des anticorps monoclonaux en thérapie a débuté en 1981 dans le cas de rejets aigus de transplant avec **muromomab**, un anticorps monoclonal murin produit par la technique des hybridomes. Cet anticorps a été administré à des patients de manière à se fixer sur le récepteur CD3 des lymphocytes T et empêcher une réponse immunitaire²⁵. Néanmoins, des inconvénients sont apparus notamment en termes de demi-vie trop courte et d'immunogénicité. Ainsi au fur et à mesure des années, les anticorps monoclonaux ont évolué vers la chimérisation, l'humanisation puis la production d'anticorps entièrement humains. Le premier anticorps totalement humain à avoir été accepté par la FDA (Food drug agency) et l'EMA (European medicines agency) et à avoir obtenu une AMM (Autorisation de Mise sur le Marché) est l'**Adalimumab** aussi dit « Humira® » prescrit dans le cas de la polyarthrite rhumatoïde mais également d'autres maladies auto-immunes comme la spondylarthrite ankylosante ou la maladie de Crohn. De plus, les formats utilisés dépendent de la cible et de la visée diagnostique ou thérapeutique. Les anticorps qui sont utilisés en imagerie sont sous format scFv ou Fab car plus furtifs et éliminés par voie rénale. En vue d'une thérapie, il peut être important de conserver le fragment Fc car il permet des fonctions biologiques telles que l'activation du complément (CDC), la cytotoxicité à médiation cellulaire (ADCC) ou encore la phagocytose (ADCP). Les sous-classes d'IgG1 à 3 sont les mieux adaptées. Elles se fixent au récepteur FcγRIII présent à la surface de cellules effectrices, notamment les cellules NK entraînant un phénomène de cytotoxicité à médiation cellulaire ou ADCC (antibody

dependent cell mediated cytotoxicity). Une revue écrite par Frenzel et al. en 2016 fait état de tous les anticorps en cours de validation clinique de la phase 1 à la phase 3 où il est à remarquer que la quasi-totalité des anticorps sont des anticorps humains³⁴.

VII Sélection d'anticorps par la technologie du « phage display »

Plus de quarante ans après leur découverte, les anticorps sont devenus de vraies armes pour la thérapie. Plusieurs anticorps monoclonaux ont obtenus une autorisation de mise sur le marché (AMM) des agences sanitaires pour leur administration dans le cas de maladies graves. Du à l'immunogénicité des anticorps murins et chimériques, les recherches se sont tournées vers le développement d'anticorps humanisés et d'anticorps humains plus affins et ayant de meilleures fonctions effectrices. Cependant l'obtention d'anticorps humains par la technique des hybridomes à partir de cellules de patients immunisés reste très problématique³⁵. Brièvement, il s'agit d'isoler les lymphocytes B d'un patient puis de faire une stimulation *in vitro* avec l'antigène d'intérêt (la GPIIbIIIa ici) afin d'induire l'expansion clonale – antigène dépendante. Une fois activés, les lymphocytes B sont fusionnés à des myélomes pour obtenir des hybridomes exprimant des anticorps monoclonaux spécifiques de la protéine d'intérêt. De nos jours, l'évolution de la biologie moléculaire permet de mimer le système immunitaire *in vitro* et de construire des banques de fragments hyper-diversifiées avec des gènes variables humains. Ces fragments peuvent ensuite être sélectionnés et exprimés à la surface de phages filamenteux. Grâce à cette technique il est possible d'avoir accès au génotype, séquence codante de l'anticorps encapsidée dans la capsid du phage, ainsi qu'au phénotype, molécule d'anticorps exprimée à la surface du phage.

Les phages filamenteux sont utilisés pour présenter à leur surface, en fusion avec le domaine amino-terminal de leurs protéines P3 ou P8, des molécules telles que des peptides aléatoires, des fragments d'anticorps (Fab, Fv ou scFv) ou d'autres protéines. La première démonstration d'expression de peptides à la surface de phages a été faite par Georges Smith en 1985³⁶. Il faudra attendre les années 90 avec l'équipe de McCafferty pour obtenir pour la première fois l'expression d'anticorps à la surface de phages³⁷. Cette technologie a permis par la suite de sélectionner des anticorps de dans de nombreuses pathologies comme l'infection par le virus de la Dengue (contre la protéine NS1 (NonStructural protein 1) du virus de la dengue^{38,39}), la filariose lymphatique (contre l'antigène BmSPX (Brugia malayi SXP protein)⁴⁰), les maladies neurologiques (contre les transporteurs VGLUT (Vesicular Glutamate Transporter)⁴¹ ou encore dans les pathologies cardiovasculaires (contre l'apolipoprotéine A1⁴², le CD100⁴³ ou encore les LDL oxydées⁴⁴). L'un des plus grands succès recensés dans la sélection par phage display d'anticorps utilisés en thérapie est l'**Adalimumab**, plus connu sous le nom d'HUMIRA qui a été mis sur le marché par la société Cambridge Antibody Technology et les laboratoires Abbott. Cet anticorps est dirigé contre le TNF α . Il est prescrit dans le cadre du traitement contre la polyarthrite rhumatoïde, le rhumatisme psoriasique ou la maladie de Crohn. L'**Adalimumab** représente un chiffre d'affaire annuel dépassant le milliard de dollars.

VII.1 Introduction sur le phage filamenteux M13

Le phage filamenteux M13 appartient au groupe II du type *Virus*, famille des *Inoviridae*, genre *Inovirus*. Le phage M13 est non lytique et capable d'infecter les bactéries *gram(-)* du type *E.Coli*. Long de 1 μ m et de 10nm de diamètre, le phage filamenteux possède un génome de 6.5kb et un ADN simple brin circulaire codant pour 10 protéines. Cinq de ces 10 protéines forment la capsid du phage : P3, P6, P7, P8 et P9 (**Figure 36**). La protéine P8 est la protéine majoritaire avec un nombre de copies supérieur au millier. Les 4 autres protéines sont représentées à raison de 5 par phage avec la P7 et P9 d'un côté et la P3 et P6 de l'autre côté. La protéine P3 permet l'infection de la cellule hôte et l'expression de la protéine de fusion. Elle se compose de trois domaines : un domaine Cterm impliqué dans l'ancrage à une des extrémités du phage et deux domaines Nterm N1 et N2 impliqués quant à eux dans

l'attachement au pilus de *E. Coli*. Les protéines P2, P5 et 10 participent à la réplication alors que P1 et P4 forment un canal permettant la libération des particules phagiques nouvellement produites. Le phage possède un tube viral susceptible de changer de taille afin d'encapsider des brins plus ou moins longs d'ADN. Ceci en fait donc un outil de choix pour le phage display. D'autres études, bien que moins nombreuses, se sont portées sur les phages fd et fl (proche du M13) ainsi que T4 et T7. Cependant un des problèmes majeurs à l'utilisation des phages T4 et T7 est la lyse bactérienne engendrée lors de la production des particules phagiques. Le phage M13 n'induisant qu'un ralentissement de la croissance bactérienne est utilisé pour la quasi-totalité des études en phage. De plus, le passage dans l'espace périplasmique des protéines produites par le phage M13 permet un meilleur repliement surtout quand la protéine requiert la présence de ponts disulfures.

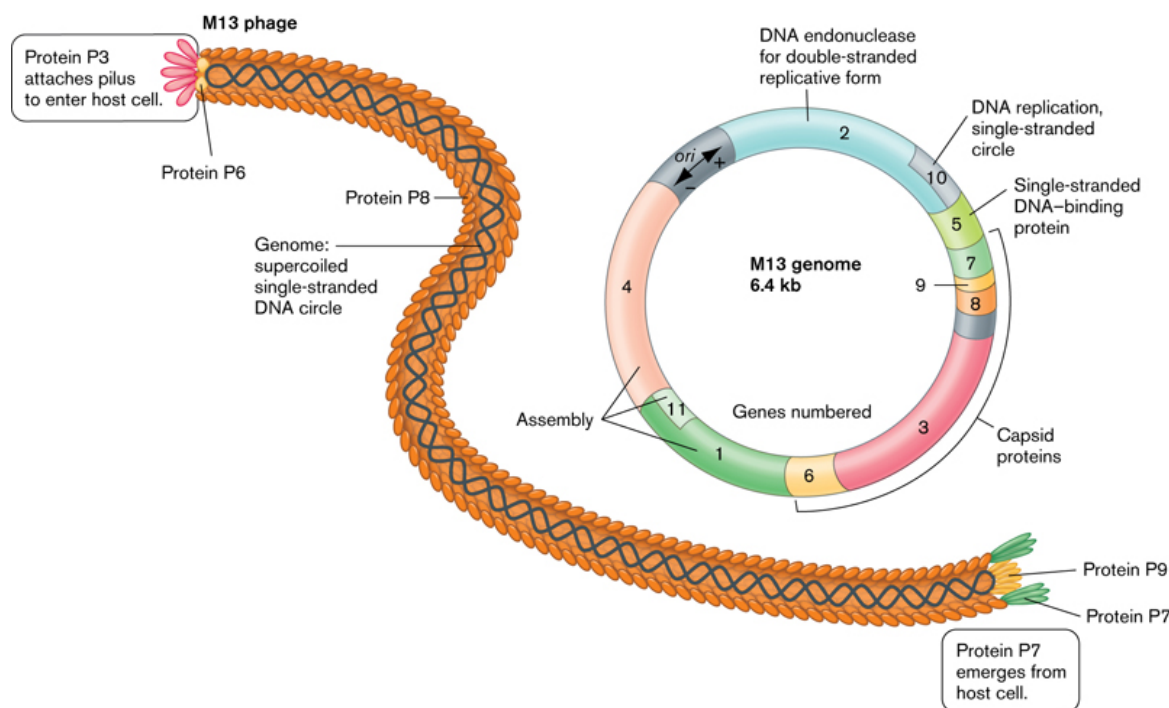


Figure 36 La structure du phage M13 et de son génome. Présentation de la structure du phage M13 avec ses protéines encodées sur le génome du phage (code couleur structure-génome). (Extrait du site <http://www.wwnorton.com/college/biology/microbiology2/ch/11/etopics.aspx>)

VII.2 Cycle d'amplification du phage M13

L'ADN du phage code pour onze gènes au total : 5 gènes codent pour des protéines qui constituent la capsidie et les autres gènes pour des protéines impliquées dans la réplication et l'assemblage de la particule virale. Lors d'une infection, l'ADN simple brin circulaire va être converti en double brin. Ce premier ADN va servir en premier lieu pour la production des protéines, en particulier de la polymérase virale. Cette polymérase va permettre la synthèse de l'ADN simple brin par un mécanisme de réplication en « rolling circle ». De cette manière, les nouvelles particules virales synthétisées s'accumulent dans le cytoplasme de la cellule bactérienne. Les protéines sont insérées dans la membrane interne, la partie Nterm dans le périplasmique et la partie Cterm dans le cytoplasme (**Figure 37**). Les protéines P4 et P1 synthétisées forment un pore entre la membrane interne et la membrane externe par lequel l'ADN viral est expulsé. L'assemblage du phage débute par l'ajout des protéines P7 et P9 à l'extrémité de la particule. S'ensuit ensuite la formation du tube viral par un enchaînement de protéines P8 englobant l'ADN simple brin. L'assemblage se termine par l'incorporation de la P3 et P6 à l'autre extrémité du tube viral. La nouvelle particule synthétisée et assemblée est alors expulsée de la cellule bactérienne sans lyse cellulaire (**Figure 37**).

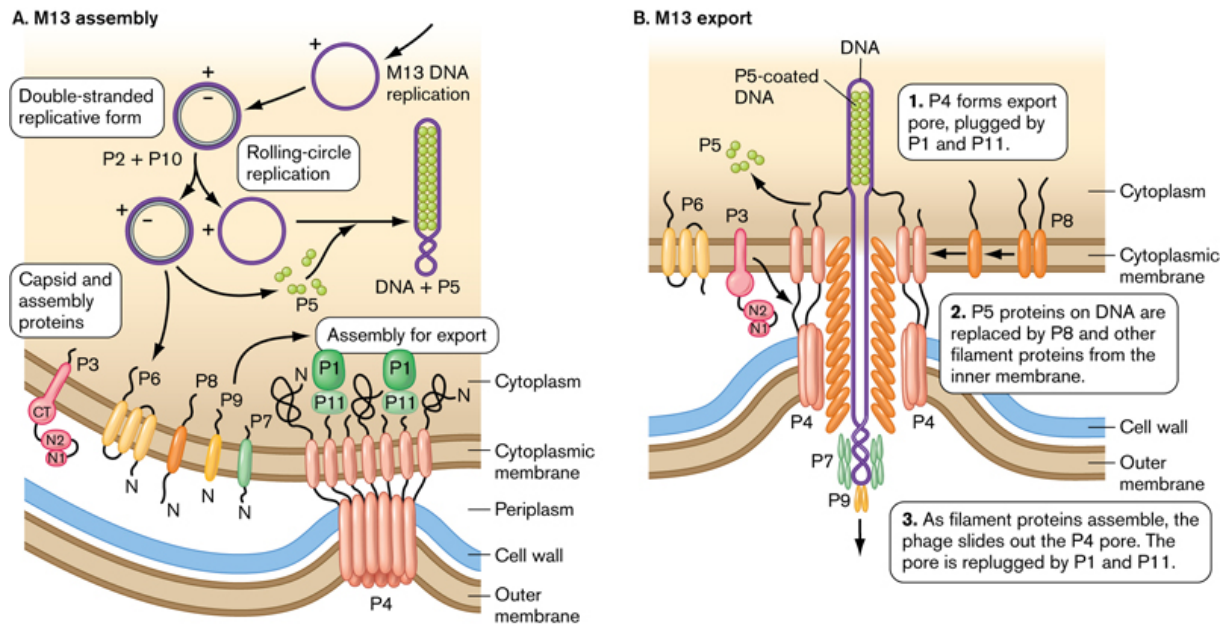


Figure 37 Assemblage et exportation de phages M13. (A) Réplication du phage M13. L'assemblage des nouveaux phages est effectué via les protéines du phage auxiliaire au niveau de la membrane cellulaire. (B) Assemblage des phages puis expulsion de ces derniers à travers le pore membranaire formé par des protéines P4. La protéine P4 forme un canal utilisé pour l'expulsion des phages filamenteux sans lyser la bactérie hôte. L'ADN monocaténaire du virus filamenteux est entouré par des protéines P8 au fur et à mesure qu'il traverse la membrane, puis il passe à travers le canal P4 pour être sécrété en tant que phage infectieux dit « mature» (Extrait du site <http://www.wwnorton.com/college/biology/microbiology2/ch11/etopics.aspx>)

VII.3 Exposition des protéines de fusion à la surface du phage

Lors de la conception des banques, les études doivent prendre en compte la protéine de surface à utiliser pour exposer la protéine de fusion. En pratique toutes les protéines de surface peuvent être utilisées à cet effet, cependant les protéines P3 et P8 semblent privilégiées. La protéine à exposer est fusionnée par la partie Nterm entre le peptide signal et la partie NH₂ de la protéine présentatrice mature. La protéine P8 est la plus représentée à raison de 2700 copies par phage. Cependant seuls des peptides de courte taille (6 à 8 acides aminés) peuvent être présentés en raison de la restriction de taille imposée par les contraintes géométriques d'assemblage des sous-unités P8 pour former la capsid. La protéine P8 de par son nombre de copies peut également permettre de présenter en grand nombre un peptide de faible affinité en augmentant l'avidité. De manière à pouvoir exprimer de plus gros fragments, comme les banques d'anticorps, la protéine P3 est la protéine de choix grâce à un nombre de copies limité à 5. Ainsi, elle permet d'exposer des polypeptides codés par des ADNc, des peptides aléatoires ou encore des fragments d'anticorps « simples chaînes » (Fv, scFv) ou des fragments Fab.

Pour la présentation et la production des particules phagiques recombinantes, deux vecteurs sont essentiels : le phagemide et le phage auxiliaire M13.

VII.4 Les phagemides

Un phagemide correspond à un plasmide auquel a été ajouté une origine de réplication d'un phage dite « M13 ori » en plus de son origine de réplication bactérienne nommée « ColE1 ori » (Figure 38). Le gène de la protéine de fusion P3-scFv est sous le contrôle d'un promoteur bactérien régulé par l'opéron lactose. De plus, pour faciliter l'étude des phages recombinants par différentes techniques comme l'ELISA (Enzyme linked immuno assay) et l'immunohistochimie après production, des tags peuvent être ajoutés comme le tag c-myc ou

le tag 6-HIS. Ces tags sont très utiles également pour la production des scFv solubles (non fusionnés à la P3). Dans la bactérie, le phagemide contenant la protéine de fusion sera préférentiellement amplifié et encapsidé dans les nouvelles particules virales. En effet, l'utilisation d'un phage auxiliaire contenant une origine de réplication défaillante permet d'apporter les protéines nécessaires à la production des nouveaux phages recombinants sans formation de phages auxiliaires natifs. Par contre, la particule virale synthétisée présentera une P3 sauvage ainsi qu'une P3 chimère. Ainsi, le phage, nouvellement produit, garde sa capacité à infecter d'autres bactéries pour la transmission de l'information génétique de la protéine d'intérêt contenue dans le phagemide encapsidé dans le phage recombinant. La conservation de l'expression de la P3 sauvage de manière simultanée à la P3 chimère permet de limiter le nombre de P3 chimère et donc les effets d'avidité conséquents non souhaités lors d'un criblage par bio-réactivité. Un des inconvénients majeurs au phage display est cependant la forte proportion de phages recombinants (c'est-à-dire contenant le phagemide) ne présentant que des P3 sauvages à leur surface. Pour contourner ce problème, Rondot et al ont développé un phage auxiliaire dit « hyperphage » ayant les mêmes caractéristiques que le phage M13 mais avec le gène codant pour la protéine P3 délété. Ainsi, la bactérie est contrainte d'utiliser le phagemide en tant que seule source de protéines P3. Le seul inconvénient de l'utilisation de ce phage auxiliaire est qu'il ne peut être amplifié dû à l'absence de ce gène requis pour l'infection ⁴⁵. Le principe de fonctionnement de l'hyperphage est présenté en **Figure 39**.

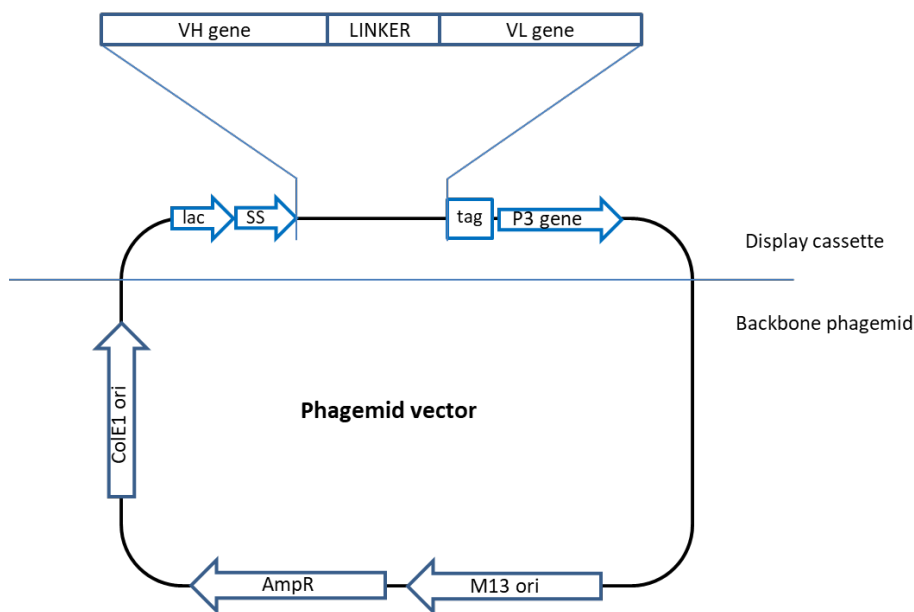


Figure 38 Vecteur phagemide. Les séquences nucléotidiques sont insérées entre deux sites de restriction. ColE1 ori : origine de réplication bactérienne ; M13 ori : origine de réplication phagique ; AmpR : gène de résistance à l'ampicilline ; lac : opéron lactose ; SS : séquence signal

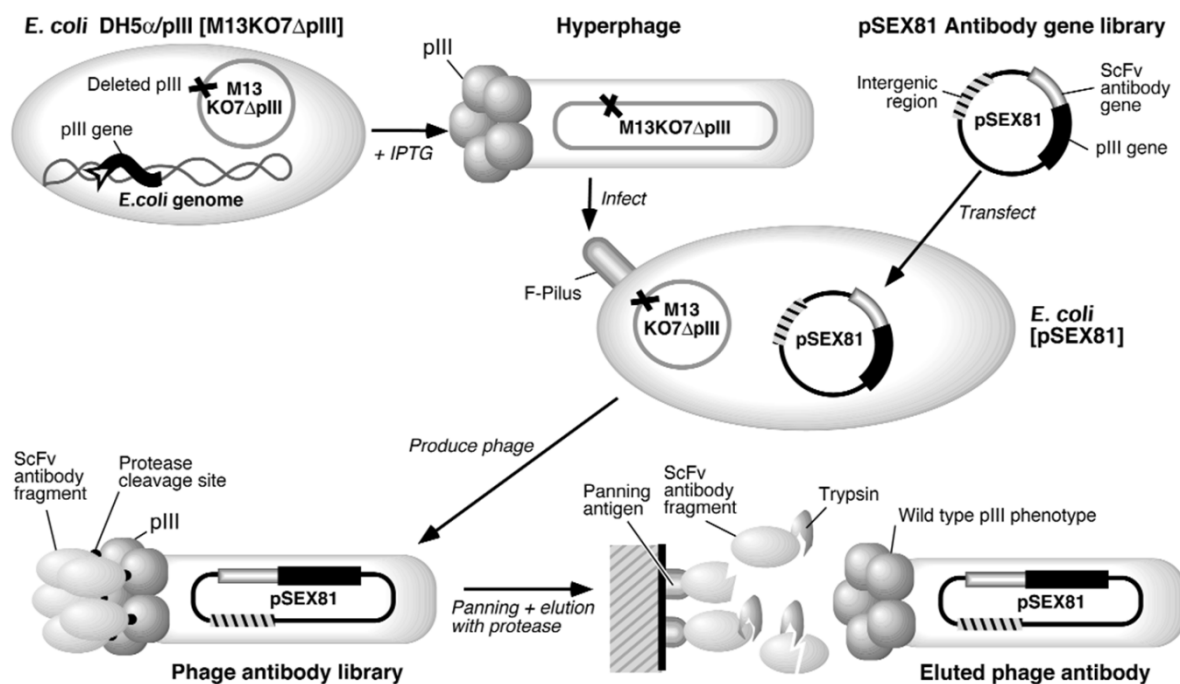


Figure 39 Le concept de l'hyperphage: schématisation du phage auxiliaire déplété du gène de P3 dit « hyperphage » suivi de l'infection classique d'une bactérie. La synthèse de l'hyperphage se fait grâce à la présence de la protéine P3 dans le génome de la bactérie. Lorsqu'il est utilisé pour l'assemblage du phagemide (pSEX ici), le phage produit exprime plusieurs copies de la protéine de fusion P3-scFv. Le site de clivage de la trypsine présent entre le scFv et la protéine P3 sert lors de l'éluion après la sélection. Il permet donc de pouvoir utiliser des tampons physiologiques sans dégrader ni le scFv ni la protéine P3 pour une ré-infection bactérienne (Extrait de l'article de Rondot et al. paru en 2000 dans *Nature Biotechnologies*).

VII.5 Les banques combinatoires d'anticorps

VII.5.1 Construction d'une banque combinatoire

Dans l'optique de la construction d'une banque combinatoire de fragments d'anticorps, l'ensemble des gènes variables exprimés du répertoire immunitaire d'un animal ou d'un individu est amplifié à partir de l'ARN messager provenant des cellules B de la rate ou du sang périphérique à l'aide d'oligonucléotides définis selon les gènes variables VH et VL du répertoire germinale^{46,47,48,49}. Ces oligonucléotides s'hybrident en 5' de la VH au niveau du FR1 ou de l'exon Leader et en 3' de JH ou JL. L'amplification des gènes VH et VL s'effectue de manière séparée afin d'obtenir deux répertoires : le répertoire des gènes VH et le répertoire des gènes VL. Il y a ensuite un assemblage aléatoire des VH et VL grâce à l'ajout d'un linker (courte séquence peptidique) s'hybridant en 3' du gène VH et en 5' du gène VL. Cette courte séquence sert de bras flexible entre un gène VH et un gène VL. En vue du clonage dans un phagemide, des sites de restriction sont insérés par l'utilisation d'amorces spécifiques lors de l'amplification de l'insert. Après un clonage en masse des inserts dans un vecteur de type phagemide, la taille du répertoire dépendra uniquement de l'efficacité de transfection dans des bactéries électro-compétentes après électroporation. Cette diversité peut varier entre 10^7 et 10^{10} fragments différents.

VII.5.2 Les banques synthétiques et semi-synthétiques

Les anticorps possèdent la capacité à reconnaître de manière spécifique un antigène. Cette reconnaissance est en particulier liée aux CDRs. Le CDR3 est le plus long et le plus variable ce qui en fait la zone centrale du paratope. Dans les années 90, les premières banques semi-synthétiques consistaient en l'introduction de CDR3H aléatoires au sein de segments VH⁵⁰. Les banques ont ensuite évolué avec l'introduction de CDR3L dans des segments VL⁵¹. Des banques synthétiques ont également été réalisées. Pour la construction de ces banques, des mutations ont été introduites dans les 6 CDRs de manière aléatoire par utilisation d'oligonucléotides chevauchants permettant par PCR d'amplifier les gènes entiers⁵². Ces mutations aléatoires ayant des effets négatifs sur la structure des anticorps, une optimisation de la construction de ces banques synthétiques a été nécessaire. Ainsi, une conservation des résidus canoniques de chaque CDR a permis de limiter ces effets négatifs^{53,54,55}. Par la suite, les banques ont évolué en vue d'optimiser les anticorps en termes d'efficacité. Pour cela, les régions CDRs aléatoires ont été insérées dans des charpentes d'anticorps connues pour leurs propriétés thermodynamiques ainsi que leur forte représentation lors d'une réponse immunitaire, comme les régions germinales DP-47 et DP-3⁵⁶. Ainsi, cette optimisation amène à la construction de banques présentant un taux de production accru et des propriétés d'agrégation moindres.

En 2007-2008, le Pr MARTINEAU a développé une nouvelle banque d'anticorps semi-synthétique entièrement humaine se basant sur une charpente d'un scFv humain nommé 13R4 et l'utilisation de CDR3 H et L semi-aléatoires, c'est-à-dire respectant la position des acides aminés retrouvés dans les régions hypervariables chez l'homme (Martineau et al. EP2134841 B1. [WO2008110914A2](#)). La charpente 13R4 présentait un avantage majeur qui est sa conformation tridimensionnelle optimale dans le cytoplasme des cellules mammifères et bactériennes favorisant ainsi la production d'anticorps fonctionnels⁵⁷. Plusieurs répertoires humains VH et VL ont été utilisés pour la construction des oligonucléotides dégénérés de manière à respecter la position de certains acides aminés retrouvés dans les CDR3 humains.

VII.5.3 Les banques naïves

Dans l'optique de la construction de ces banques, les gènes variables VH et VL sont issus de l'ARNm codant pour des IgM provenant de cellules lymphocytaires B issues de patients non immunisés^{58, 59}. Ces banques présentent l'avantage d'un large choix d'anticorps. Cependant deux inconvénients prédominent : en premier lieu une faible affinité comme observée lors d'une primo réponse ($10^{-6} - 10^{-7}$ M) ; en deuxième lieu, ces banques ne permettent pas de sélectionner des anticorps spécifiques de molécules du soi. Ceci peut être un réel problème dans le cas de recherches d'anticorps thérapeutiques dans le cadre de maladies auto-immunes telles que la maladie de Crohn, le purpura rhumatoïde ou encore la polyarthrite rhumatoïde et la sclérose en plaque.

VII.5.4 Les banques immunes

Contrairement aux banques naïves, les gènes VH et VL utilisés pour la construction de ce type de banques sont issus de l'ARNm codant pour les IgG de cellules lymphocytaires B de patients immunisés. Dans le respect des règles d'éthique, il est difficile d'envisager une immunisation du patient avec un antigène donné en vue d'obtenir des anticorps spécifiques de ce même antigène. Ces banques sont donc utilisées à des fins de sélection contre des antigènes d'intérêt dans le cadre d'une pathologie donnée. Le cas des maladies auto-immunes est particulièrement intéressant ici pour l'étude du répertoire immunitaire des patients^{60,61}. Ce type de banque a permis d'isoler des anticorps humains de très haute affinité ($10^{-8} - 10^{-9}$ M).

Le Pr MONDON a ainsi créé en 2006 une banque hyper-diversifiée à partir d'ARNm codant pour les IgG provenant de lymphocytes B de personnes saines et de patients atteints de diverses pathologies telles que les cancers, les maladies auto-immunes et les maladies

inflammatoires. A partir des ARNm, ces derniers ont été rétro-transcrits en ADNc. Afin de reproduire le phénomène d'hypermutation somatique décrit précédemment, des PCR ont été réalisées avec des polymérases dites de faible fidélité. Les deux répertoires VH et VL résultants ont ensuite suivi le processus de construction de banque combinatoire présenté dans la section VI.6.1.

VII.6 Phage display *in vitro*

Afin de pouvoir sélectionner des anticorps spécifiques d'antigènes, le phage display *in vitro* est la méthode de choix actuellement la plus utilisée en laboratoire. Cette technique consiste à présenter des phages recombinants à des antigènes identifiés, purifiés ou recombinants, fixés sur un support solide. Cela peut devenir plus complexe lorsque l'objectif est de sélectionner un anticorps sur un récepteur exprimé par une cellule ou sur un des protéines membranaires.

VII.6.1 Sélection sur un antigène isolé

Il s'agit ici de la technique la plus utilisée en laboratoire à l'heure actuelle. L'antigène « connu » préalablement immobilisé sur une plaque est mis en présence de phages recombinants exprimant des anticorps ou de peptides. Les phages exprimant une molécule spécifique de cet antigène vont se fixer (**Figure 40**). Après de multiples lavages, une élution acide ou en présence de trypsine (si un site de clivage entre la P3 du phage et la molécule exprimée est présent) est effectuée. Afin d'affiner la sélection, les phages élués peuvent être amplifiés dans la bactérie afin de réaliser des tours successifs de sélection. Il est à prendre en compte que ce type de sélection peut aboutir à des anticorps qui une fois injectés seront dans l'incapacité de reconnaître leur cible du fait qu'ils aient été préalablement sélectionnés sur des antigènes fixés pouvant avoir subi des modifications conformationnelles. Afin de pallier ce problème, les antigènes peuvent être mis en solution après un marquage à la biotine⁶² (**Figure 41**).

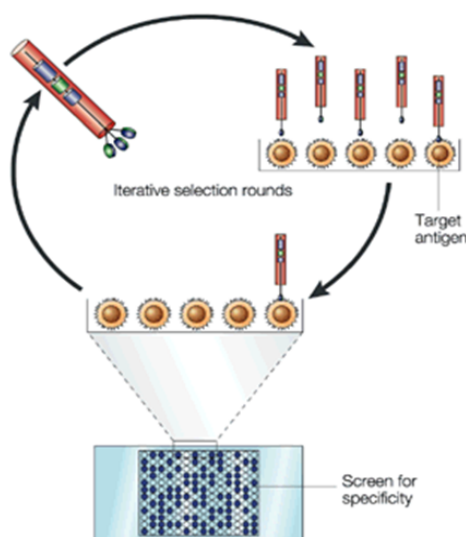


Figure 40 Phage display *in vitro* sur un antigène isolé et immobilisé. Le processus *in vitro* est basé sur la sélection de banques d'anticorps contre une cible immobilisée. Les phage-anticorps non liés sont éliminés par lavage et les anticorps récupérés sont amplifiés par infection de bactéries Escherichia coli. Les cycles de sélection sont ensuite répétés pour obtenir un enrichissement en anticorps spécifiques. Le format d'anticorps pour la sélection est soit un Fab soit un scFv. Le phage display a permis de sélectionner des dizaines de milliers de clones spécifiques. Les fragments d'anticorps eux-mêmes peuvent être utilisés comme agents thérapeutiques mais ils peuvent également être transformés en immunoglobulines par le clonage des gènes variables dans un plasmide en incorporant les gènes de région constante des immunoglobulines. Les gènes sont transfectés dans les lignées cellulaires qui produisent donc des immunoglobulines entièrement humaines. (Extrait de l'article de Henrik Breker et Inger Sandlie, 2002)

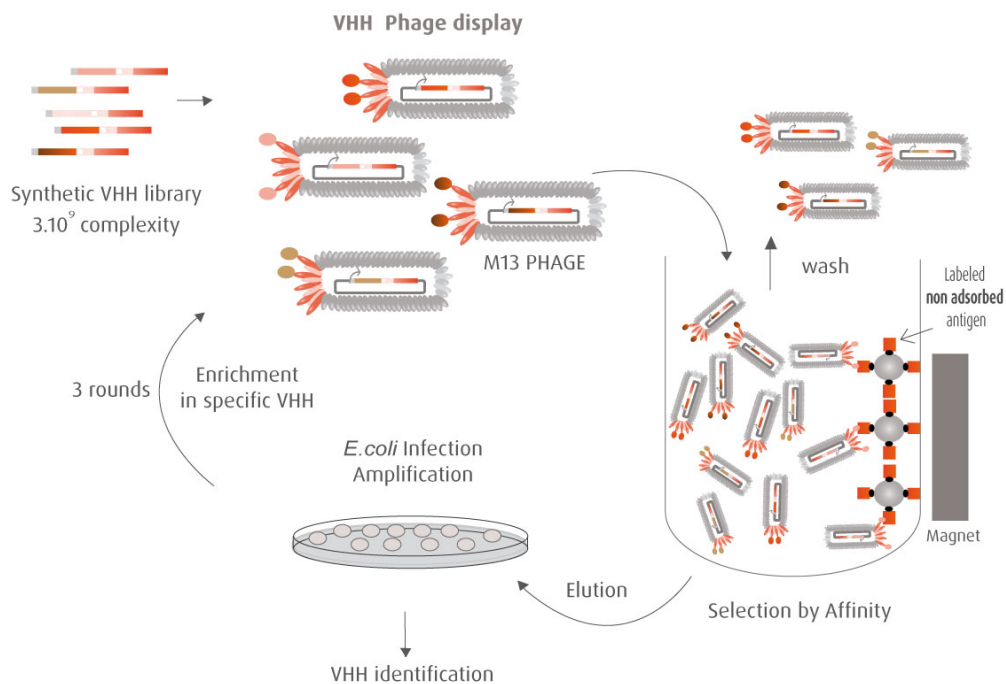


Figure 41 Phage display *in vitro* sur un antigène biotinylé adsorbé sur des billes/streptavidine. La société Hybrigenics propose la sélection d'anticorps mono-chaînes contre des antigènes en solution. Pour la sélection de phage-anticorps, les phages présentant des anticorps de la librairie hs2dAb contenant $3 \cdot 10^9$ V_HH sont mis en contact avec l'antigène d'intérêt adsorbé sur des billes. Après une étape de lavage, des anticorps sélectionnés reconnaissant l'antigène sont élués puis amplifiés dans des bactéries. Par la suite, un autre cycle de sélection de phage-anticorps peut commencer (Extrait de <https://www.hyribody.com/contenu/ab-selection-technologies-menu/phage-display-antibody-selection>).

VII.6.2 Sélection sur cellules

Parmi les protéines pouvant être des cibles thérapeutiques, les protéines membranaires représentent un fort pourcentage. Cependant, ces dernières posent un réel problème d'extraction car souvent endommagées par les tampons. Ainsi pour contourner ce problème, des équipes ont envisagé la sélection de banques combinatoires directement sur des cellules exprimant ces récepteurs^{63,64,65} (**Figure 42**). Ce type de sélection comporte tout de même des inconvénients notamment la fixation non spécifique des banques sur la membrane. Pour pallier cette éventualité, des équipes de recherche utilisent la contre sélection sur des surfaces membranaires^{64,65} ou encore l'éluion par action de compétiteurs des récepteurs membranaires⁶⁶.

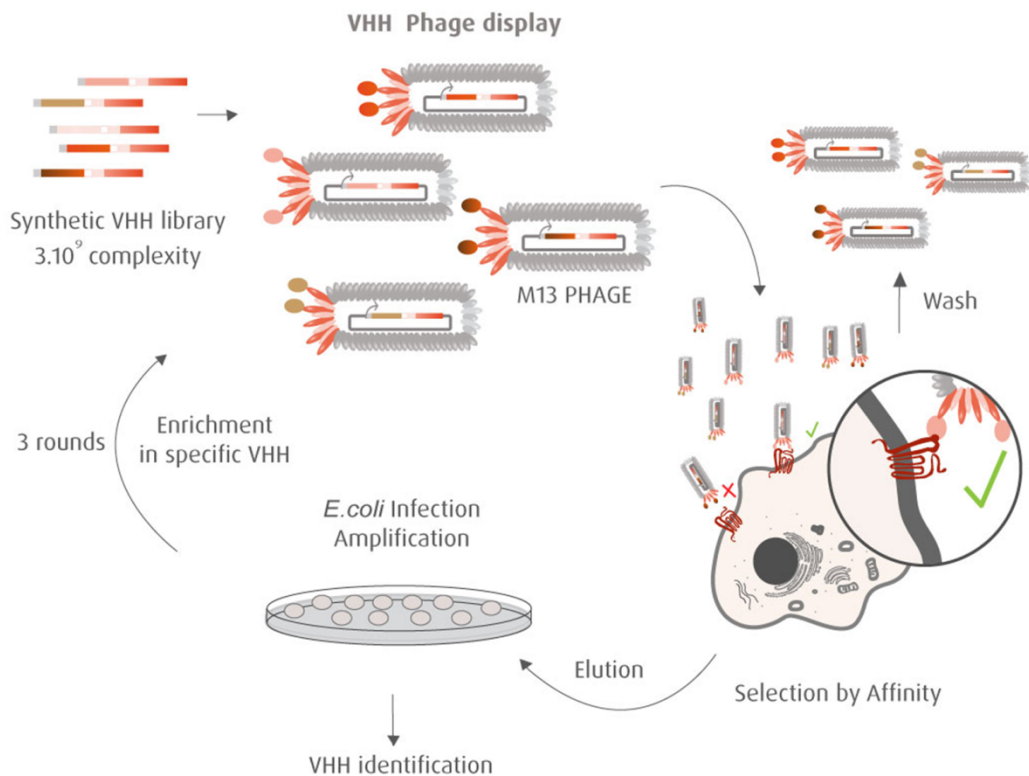


Figure 42 Phage display sur des cellules exprimant le récepteur d'intérêt. La société Hybrigenics propose de sélectionner des anticorps par Phage Display directement sur des cellules vivantes lorsque la sélection *in vitro* n'est pas possible, par exemple pour des récepteurs avec des domaines transmembranaires comme les GPCR. La librairie hs2dAb contenant 3.10^9 V_HH est exprimée à la surface de phages et incubée avec la cellule exprimant l'antigène d'intérêt. Après une étape de lavage avec des cellules qui n'expriment pas l'antigène, l'éluion des anticorps spécifiques est effectuée et leur enrichissement est déterminé (Extrait de <https://www.hyribody.com/contenu/ab-selection-technologies-menu/phage-display-antibody-selection>).

VII.6.3 Sélection sur sérum de patients

Dans le cadre de maladies auto-immunes ou inflammatoires, le patient développe des anticorps dirigés contre des cibles pouvant être intéressantes dans l'optique du développement d'un traitement. Pour cela, de nouvelles méthodes de sélection ont été développées. Il s'agit de sélectionner des banques de peptides contre le sérum de patients afin de mimer l'épitope reconnu par l'anticorps. Ainsi, ces peptides peuvent se fixer de manière spécifique sur des anticorps de patients malades mais pas sur ceux d'un individu sain. Ainsi, il a été possible de sélectionner des peptides mimant des épitopes antigéniques reconnus par des anticorps synthétisés par des patients infectés par le virus de l'Epstein-Barr⁶⁷, des patients ayant eu un traumatisme intracrânien⁶⁸, des patients atteints de sclérose en plaque⁶⁹, des patients atteints de l'hépatite B⁷⁰.

VII.6.4 Conclusions

Les méthodes citées ci-dessus ne sont pas exhaustives. D'autres techniques de sélection existent. Cependant, l'étape de confirmation des cibles par injection *in vivo* d'anticorps obtenus par phage display *in vitro* aboutit souvent à une absence de reconnaissance de la cible. En effet, lors de l'extraction des cibles, ces dernières peuvent perdre leur conformation initiale. De plus, l'environnement influence les protéines, ce qui est un facteur important à

prendre en compte. De manière à obtenir des anticorps reconnaissant leur cible dans leur conformation native, les techniques de sélection ont évolué vers le phage display *in vivo*.

VII.7 Phage display *in vivo*

Le phage display *in vivo* est une technique permettant d'obtenir des ligands spécifiques de différentes surfaces endothéliales, chacune étant différente suivant la pathologie étudiée. En effet, l'endothélium exprime des molécules spécifiques en fonction de sa localisation tissulaire et, comme nous l'avons vu précédemment, en fonction de son état pathologique. Comme l'on décrit Garlanda et Ruoslahti, chaque endothélium possède une signature bien particulière. Dans le cadre de l'athérosclérose, l'endothélium va exprimer des molécules d'adhésion pour le recrutement de cellules^{71,72}. De ce fait ces molécules peuvent être choisies comme cibles pour de la thérapie et/ou de l'imagerie. Cependant comme l'a démontré Augustin et son équipe en 1994, les cellules endothéliales sorties de leur microenvironnement perdent leur intégrité et leur fonction⁷³. Pour pallier cela, de nombreuses techniques ont été développées notamment les co-cultures de cellules. Cependant, cela ne permet pas de reproduire totalement l'environnement de ces cellules tel qu'il est dans l'organisme. Les premières études de phage display *in vivo* sont apparues en 1996 avec l'étude de Pasqualini *et al.*⁷⁴. Cette étude a permis d'identifier plusieurs ligands au niveau des cellules endothéliales du foie et des reins après injection *in vivo* chez une souris d'une banque de peptides. Lors de cette sélection, plusieurs tours ont été réalisés, conduisant à une augmentation de l'affinité et un enrichissement des phages spécifiques exprimant à leur surface des peptides. Ainsi, ce type de sélection permet d'effectuer une sélection négative et une sélection positive de manière simultanée. Après injection, les phages/peptides vont se fixer sur les récepteurs endothéliaux dans la circulation sanguine : c'est la sélection négative. Seuls ceux fixés sur les organes d'intérêt seront récupérés : c'est la sélection positive. D'autres études ont également été conduites *in vivo* chez un animal (rat ou souris) amenant à l'identification de cibles dans le cadre du carcinome rénal⁷⁵, du carcinome hépatocellulaire⁷⁶, du cancer buccal^{77,78}. Comme illustré dans la **Figure 43**, la première étude de phage display *in vivo* chez l'homme a été faite en 2001 par l'équipe de Pasqualini *et al* après injection *in vivo* chez un patient dans un état comateux d'une banque de peptides. Ainsi, un peptide spécifique du récepteur à l'IL-11 exprimé au niveau des tumeurs de la prostate a été découvert^{79,80,81}. Cette cible sera confirmée quelques années plus tard par Zurita⁸². Grâce à ce peptide, il a été démontré que l'IL-11R α était une cible potentielle pour le traitement du cancer de la prostate. Le peptide anti-récepteur à l'IL-11 a été couplé à un peptide pro-apoptotique pour détecter l'apoptose des cellules tumorales après internalisation. Ce même peptide a été par ailleurs couplé à une sonde fluorescente, travail qui a permis de visualiser l'accumulation de la sonde dans les tumeurs du tibia de trois souris ayant subi des xénogreffes avec des cellules tumorales de la prostate⁸³. Staquicini, en 2011, a injecté la banque de peptides récupérée après l'injection chez le premier patient pour l'injecter à deux autres patients de manière synchrone (patient 1 atteint d'un lymphome non-hodgkinien ; patient 2 atteint d'un cancer de la prostate avec des métastases osseuses). Après un séquençage des phages récupérés au niveau de la prostate, du foie, de tumeurs métastatiques, de la peau, du tissu adipeux et des muscles squelettiques, une totalité de 15 autres couples peptide/ligand potentiels ont été isolés⁸⁴. Quatre de ces couples ont été confirmés expérimentalement et ont amené à la découverte de quatre ligand/cible : l'intégrine $\alpha 4$ se liant à l'annexine A4, la cathepsine B à l'apolipoprotéine E3, l'annexine A2 à la prohibitine et le récepteur des AGE à la protéinase 3 des leucocytes. Un des nombreux avantages de ligands identifiés par phage display *in vivo* est qu'ils ont été sélectionnés pour leur affinité vis-à-vis d'une cible dans des conditions physiopathologiques et sont donc des ligands incontournables pour une imagerie et une thérapie ciblée. De nombreuses études ont montré une augmentation de l'effet d'une drogue après couplage des peptides identifiés par phage display *in vivo* à des cargos chargés en molécules chimiques (liposomes chargés en doxorubicine et couplés au peptide découvert dans le cadre du cancer buccal⁷⁷).

L'athérogenèse faisant intervenir de nombreux protagonistes moléculaires et cellulaires en fait une pathologie intéressante pour une étude par phage display *in vivo*. Dans le cadre des maladies cardiovasculaires, un peptide découvert *in vivo* homologue à la protéine VLA-4 a été couplé à des nanoparticules et a montré, après injection à des souris ApoE^{-/-}, un ciblage des VCAMs exprimés à la surface des cellules endothéliales⁸⁵. Le but de notre étude qui sera décrite en section « Résultats » a pour but l'identification de nouvelles cibles grâce à la découverte d'anticorps humains capables de cibler de manière spécifique l'athérome après injection dans un lapin sous régime hypercholestérolémiant. Les prémisses de cette étude ont été réalisés en 2006 et 2010 par deux doctorants ayant développé des anticorps susceptibles de cibler la plaque avec la découverte de nouveaux biomarqueurs comme la carbonic anhydrase 2⁸⁶ et la vitronectine⁸⁷. Ainsi, dans une perspective proche, ces nouveaux anticorps pourront être utilisés en tant qu'agents ciblant après couplage à des nanoparticules pouvant être chargées en drogues mais également être étudiés pour leurs rôles antagonistes potentiels dans le processus d'athérogenèse.

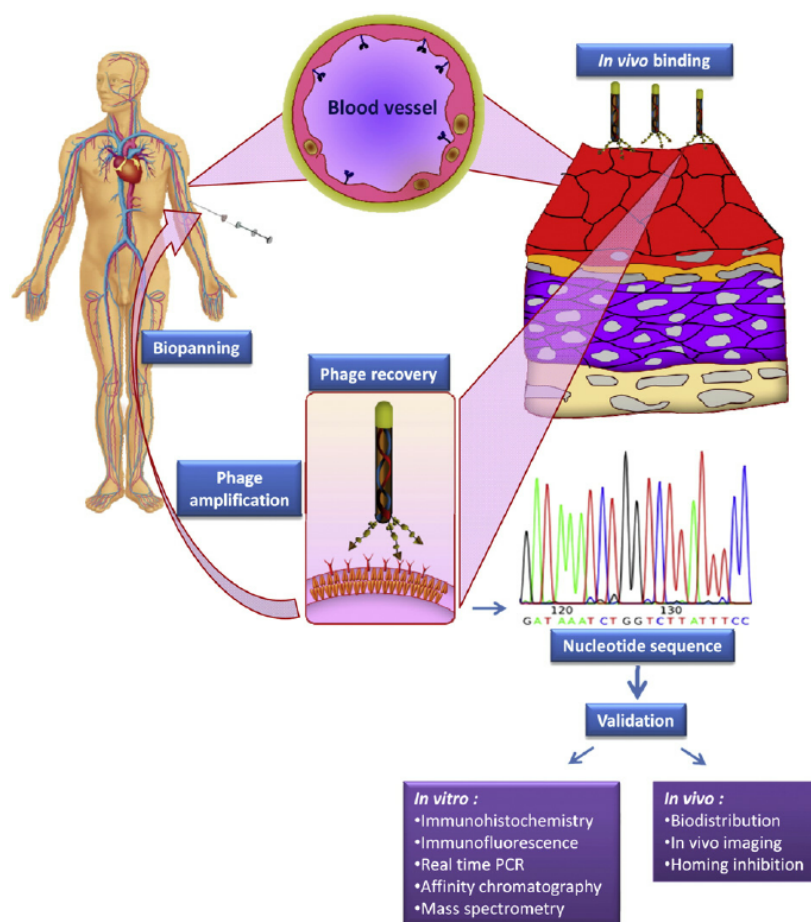


Figure 43 Phage display *in vivo* chez l'Homme. Étapes de la technologie de sélection de phage display *in vivo* pour cartographier la vasculature humaine. Les bibliothèques de phage sont injectées par voie intraveineuse pour identifier les phage-peptides qui interagissent avec des molécules spécifiquement exprimées dans les vaisseaux sanguins d'un organe ou d'un tissu particulier. Après l'élimination des organes cibles ou des biopsies tissulaires, les phages sont récupérés et amplifiés afin d'effectuer un second tour de sélection. Après identification des séquences, elles sont traduites en séquence protéique. Une validation soit *in vitro* (immunohistochimie, immunofluorescence) soit *in vivo* (biodistribution, imagerie *in vivo*, inhibition spécifique par ligand compétitif) de ces peptides synthétisés est réalisée (Extrait de l'article de D'Onofrio et al. paru en 2014 dans *Biochim Biophys Acta*.)

VIII Séquençage des anticorps

La sélection par phage display de fragment d'anticorps Fab ou scFv est très utilisée en vue de sélectionner des anticorps spécifiques et d'identifier de nouvelles cibles pour le diagnostic et la thérapie dans le cadre de nombreuses pathologies comme le cancer, les maladies auto-immunes et les maladies infectieuses (**Figure 44**). Jusqu'à aujourd'hui, les anticorps ou peptides étaient criblés par des techniques *in vitro* : ELISA ou cytométrie de flux suivies d'un séquençage Sanger (séquençage de première génération) puis d'immunoprécipitation et spectrométrie de masse pour la détermination de la cible. Cependant, ces techniques de sélection sont coûteuses en temps et ne permettent de cribler les clones que par centaines. Pour pallier ce problème, le séquençage de dernière génération nommé « NGS » pour « Next generation sequencing » (deuxième et troisième génération) apparaît aujourd'hui comme une technique privilégiée pour une caractérisation haut-débit des immunoglobulines en vue de découvrir de nouveaux ligands pour l'imagerie et la thérapie (**Figure 44**).

Le premier tableau ci-dessous (**Figure 45**) détaille les avantages et les inconvénients de chaque génération de séquençage. Il est à retenir que la technologie a fondamentalement changé. L'évolution a permis de passer d'une analyse classique Sanger à une analyse directe de la molécule en temps réel. En termes de résolution, la troisième génération permet de pouvoir analyser une molécule par une molécule. De plus, la troisième génération de séquençage permet de lire des molécules beaucoup plus longues qu'auparavant allant bien au-delà du kpb (Pacific Biosciences peut aller jusqu'à 25kpb environ). Au fur et à mesure de l'évolution du séquençage, le nombre de données récoltées étant en large expansion, les techniques d'analyses informatiques ont évolué également afin de pouvoir les traiter. La première génération se réfère au séquençage Sanger, quant à la deuxième génération (SGS = Second Generation Sequencing), il s'agit des méthodes comme celles proposées par Illumina, Roche 454 ou Ion torrent pour les plus connues. La troisième génération (TGS = Third Generation Sequencing) fait référence à la technologie PacBio et Oxford nanopore. Le tableau d'après (**Figure 46**) reprend point par point les techniques de séquençage deuxième et troisième générations utilisées à l'heure actuelle avec les limitations de chacune. De nos jours, les études de séquençage des immunoglobulines utilisant les techniques, Roche 454, Illumina et Ion torrent sont les plus présentes.

A l'heure actuelle, de nombreuses études rapportent l'analyse soit du VH et du VL séparés soit uniquement VH ou de régions CDRs pour tenter de retrouver la ou les séquence(s) de clones préférentiellement enrichis au cours des tours de sélection. L'étude de Sai T Reddy, parue dans *Nature Biotechnologies* en 2010, rapporte la découverte d'anticorps spécifiques de trois antigènes à partir de l'étude du répertoire des plasmocytes, cellules dérivées de l'activation des lymphocytes B, sans passer par les méthodes de criblage⁸⁸. Les gènes variables de la chaîne lourde et les gènes variables de la chaîne légère des immunoglobulines ont été amplifiés par des amorces dégénérées puis séquencés de manière séparée par la technique Roche 454. Les données de séquençage ont été analysées par bio-informatique. La fréquence des VH et la fréquence des VL ont été déterminées. Ainsi, ces fréquences ont permis de construire des scFvs en associant le VH le plus représenté au VL le plus représenté également. L'étude de Ravn, parue dans *Nucleic acid research* en 2010, a montré deux points importants après séquençage par la technique Illumina des VH et VL séparés⁸⁹. Premièrement, par cette technique de séquençage haut débit, il est possible de déterminer l'enrichissement spécifique se produisant durant la sélection et notamment en se basant sur les CDR3. Deuxièmement, grâce à des amorces dessinées par rapport au CDR3 le plus enrichi, il est possible de faire du « rescue », c'est-à-dire d'aller repêcher des séquences entières d'anticorps. L'étude de Spiliotopoulos, parue dans *Journal of Immunological Methods* en 2015, a démontré que la séquence d'anticorps spécifiquement enrichis au cours des tours de sélection pouvait être récupérée uniquement par l'étude du CDR3H/FR4⁹⁰. Grâce à des données provenant d'un séquençage par Ion torrent des gènes variables de la chaîne lourde, des amorces basées sur le FR4/CDR3H ont été dessinées afin de récupérer les clones préférentiellement enrichis durant

la sélection. Il ne s'agit ici que de quelques exemples mais il est important de souligner qu'aucune étude n'a montré jusqu'à aujourd'hui le séquençage de régions V entières sans avoir à faire appel à des méthodes de recouplement pour rassembler les deux régions V. De plus, du fait de la haute diversité de la région V des immunoglobulines et de l'information cruciale contenue dans la séquence entière d'un scFv (association de la VH et de la VL, linker compris), le challenge reste le séquençage des régions V entières. Pour cela, les chercheurs se tournent à l'heure actuelle vers des techniques plus évoluées qui permettent d'aller au de-là du kbp séquencé, comme par exemple la technique PacBio allant jusqu'à 25kpb.

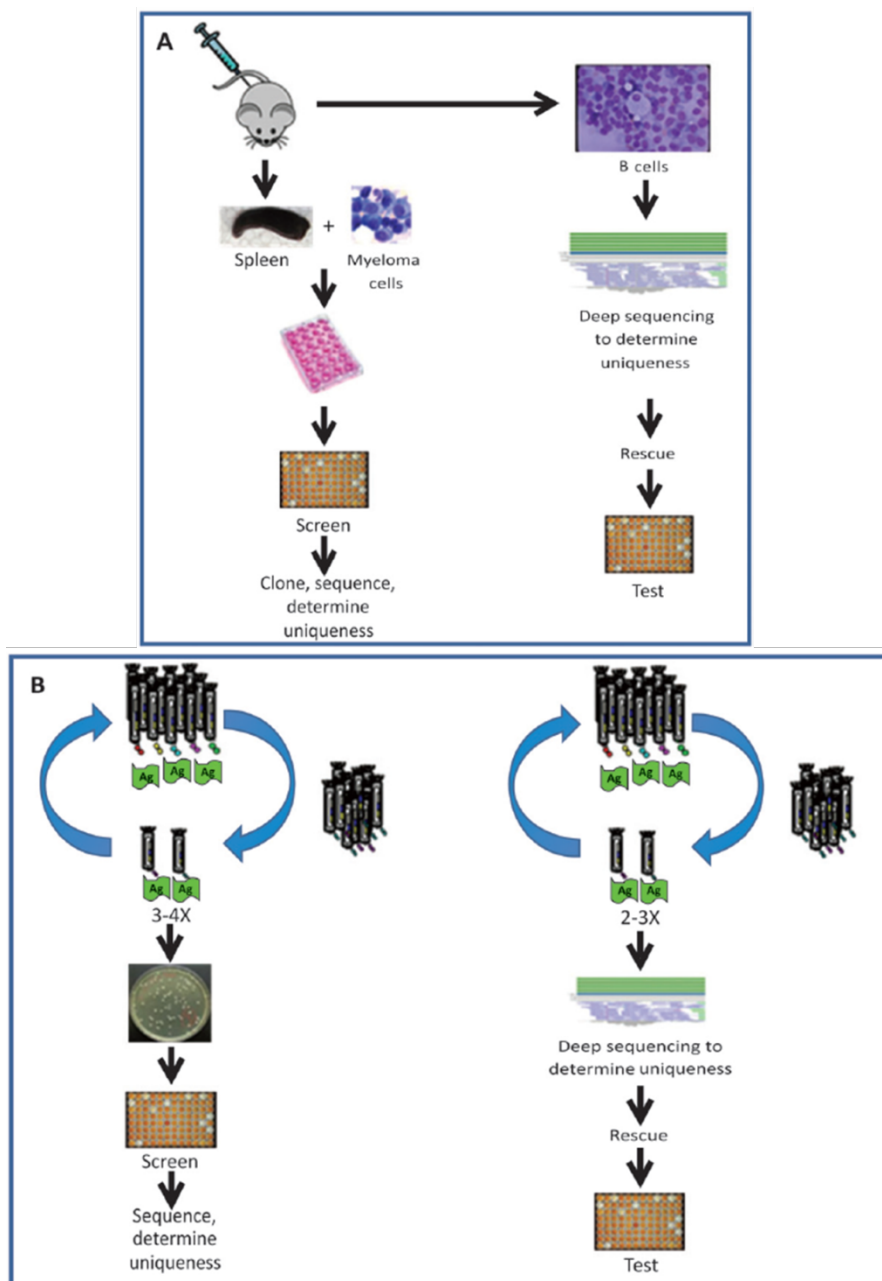


Figure 44 Comparaison de deux méthodes de travail pour la découverte d'anticorps A) dérivés de cellules B (y compris les hybridomes) et B) dérivés du phage display. Dans les méthodes de sélection traditionnelles par la technique des hybridomes et du phage display (schémas de gauche dans les encadrés A et B, respectivement), l'information sur la séquence reste indéterminée jusqu'à la fin du criblage. En conséquence, les chercheurs doivent choisir parmi les anticorps qui ont été sélectionnés par un criblage biologique. Alternativement, avec un séquençage haut débit (schémas à droite dans les encadrés A et B), les chercheurs peuvent sélectionner des clones divers pour une analyse plus poussée en début de processus (Extrait de l'article de Naso et al. paru en 2014 dans la revue *Current Drug discovery technologies*).

	First generation	Second generation ^a	Third generation ^a
Fundamental technology	Size-separation of specifically end-labeled DNA fragments, produced by SBS or degradation	Wash-and-scan SBS	SBS, by degradation, or direct physical inspection of the DNA molecule
Resolution	Averaged across many copies of the DNA molecule being sequenced	Averaged across many copies of the DNA molecule being sequenced	Single-molecule resolution
Current raw read accuracy	High	High	Moderate
Current read length	Moderate (800–1000 bp)	Short, generally much shorter than Sanger sequencing	Long, 1000 bp and longer in commercial systems
Current throughput	Low	High	Moderate
Current cost	High cost per base Low cost per run	Low cost per base High cost per run	Low-to-moderate cost per base Low cost per run
RNA-sequencing method	cDNA sequencing	cDNA sequencing	Direct RNA sequencing and cDNA sequencing
Time from start of sequencing reaction to result	Hours	Days	Hours
Sample preparation	Moderately complex, PCR amplification not required	Complex, PCR amplification required	Ranges from complex to very simple depending on technology
Data analysis	Routine	Complex because of large data volumes and because short reads complicate assembly and alignment algorithms	Complex because of large data volumes and because technologies yield new types of information and new signal processing challenges
Primary results	Base calls with quality values	Base calls with quality values	Base calls with quality values, potentially other base information such as kinetics

Figure 45 Tableau comparatif des trois générations de séquençage. (a) Il existe de nombreuses technologies TGS en développement mais peu ont été utilisées en pratique. Bien qu'il existe un potentiel important des TGS pour améliorer radicalement le débit de production des données et la longueur de lecture, les limites pratiques de ces technologies restent encore à explorer. De plus, les SGS subissent également des améliorations qui pourraient permettre d'augmenter le débit de production des données et également la longueur de lecture des brins d'ADN (Extrait de l'article de Schadt et al. paru en 2010 dans la revue *Human Molecular Genetics*)

Instrument	Length of reads	Quality	Throughput Mb/day	Turn around time	Output	Cost per run US\$	Advantages	Disadvantages	Sequencing mechanism	Instrument cost	Major used
Sanger	800	Excellent	1	2 hours	0.000,1 Gb	4/one sequence	High quality sequences, fast turn around, long reads	Low throughput, high cost	Dideoxy chain termination	95,000	Targeted sequencing
Roche 454	1000	Excellent	750	23 hours	0.7 Gb	7,000	Long reads especially useful for antibody V region sequencing, fast turn around	High error rate with polybase more than 6, high cost	Pyrosequencing, CCD light imaging	500,000 - 700,000	Targeted region sequencing, SNP detection
Ion Torrent	35-400	Good	10,000	4 hours	10 Gb	1,000	Low cost, fast turn round	High error rate	Semiconductor-based nonoptical detection, standard dNTP chemistry	50,000 -	Targeted or Exome sequencing that does not require very deep sequencing
Illumina HiSeq	75-150	Excellent	50,000	11 days	600 Gb	6,000	High throughput, widest use	Short reads, long turn around	Sequencing by synthesis, flow cell based, reversible dye termination, four color optical imaging	500,000 - 900,000	Whole genome, whole exom sequencing, SNP detection
SOLiD	50-75	Excellent	45,000	7 days	320 Gb	15,000	Accuracy, high throughput	Short reads, long turn around	Ligation and two-base coding, flow cell, four color imaging	500,000 - 700000	De Novo Sequencing, whole genome, whole exom sequencing, SNP detection
Pacific Bio	3,000	Poor	120	10 hours	0.12 Gb	500	Very long reads	High error rate	Individual polymerase, single molecule sequencing using fluorescent dNTPs	750,000	Full microbial genome sequencing, long read sequencing
Oxford Nanopore	>10, 000	Poor	>10	Flexible	Flexible	900	Longest reads, single molecule sequencing	High error rate	Single molecule pyrosequencing, uses nanopores as biosensors	No data	Full microbial genome sequencing, long read sequencing, direct RNA sequencing

Figure 46 Résumé et comparaison des nouvelles méthodes de séquençage haut débit utilisées à l'heure actuelle. (Extrait de l'article de Naso et al. paru en 2014 dans la revue *Current Drug Discovery Technologies*)

Bibliographie

1. Edward Jenner (1749-1823). Available at: <http://broughttolife.sciencemuseum.org.uk/broughttolife/people/edwardjenner>. (Accessed: 10th October 2017)
2. Pasteur, Pierre Darmon | Fayard. Available at: </pasteur-9782213594040>. (Accessed: 10th October 2017)
3. Tan, S. Y. & Dee, M. K. Elie Metchnikoff (1845-1916): discoverer of phagocytosis. *Singapore Med. J.* **50**, 456–457 (2009).
4. Behring, null & Kitasato, null. [On the development of immunity to diphtheria and tetanus in animals]. *Dtsch. Med. Wochenschr.* 1946 **90**, 2183 (1965).
5. Jules Bordet (1870-1961). *Institut Pasteur* (2016). Available at: <https://www.pasteur.fr/fr/institut-pasteur/jules-bordet-prix-nobel-1919>. (Accessed: 10th October 2017)
6. Burnet, F. M. A modification of Jerne's theory of antibody production using the concept of clonal selection. *CA. Cancer J. Clin.* **26**, 119–121 (1976).
7. Alberts, B. *et al.* The Adaptive Immune System. (2002).
8. Ribatti, D. Edelman's view on the discovery of antibodies. *Immunol. Lett.* **164**, 72–75 (2015).
9. 10 - Antibody Fc engineering for optimal antibody performance. in *Therapeutic Antibody Engineering* (eds. Strohl, W. R. & Strohl, L. M.) 225–595 (Woodhead Publishing, 2012). doi:10.1533/9781908818096.225
10. Hentges, F. B lymphocyte ontogeny and immunoglobulin production. *Clin. Exp. Immunol.* **97**, 3–9 (1994).
11. Rao, C. V. *An Introduction to Immunology*. (CRC Press, 2002).
12. Hozumi, N. & Tonegawa, S. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc. Natl. Acad. Sci. U. S. A.* **73**, 3628–3632 (1976).

13. Holliger, P. & Hudson, P. J. Engineered antibody fragments and the rise of single domains. *Nat. Biotechnol.* **23**, 1126–1136 (2005).
14. Chames, P. & Baty, D. Bispecific antibodies for cancer therapy: the light at the end of the tunnel? *mAbs* **1**, 539–547 (2009).
15. Moutel, S. *et al.* A multi-Fc-species system for recombinant antibody production. *BMC Biotechnol.* **9**, 14 (2009).
16. Shinkawa, T. *et al.* The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *J. Biol. Chem.* **278**, 3466–3473 (2003).
17. Niwa, R. *et al.* IgG subclass-independent improvement of antibody-dependent cellular cytotoxicity by fucose removal from Asn297-linked oligosaccharides. *J. Immunol. Methods* **306**, 151–160 (2005).
18. Marchal, I., Mir, A. M., Kmiécik, D., Verbert, A. & Cacan, R. Use of inhibitors to characterize intermediates in the processing of N-glycans synthesized by insect cells: a metabolic study with Sf9 cell line. *Glycobiology* **9**, 645–654 (1999).
19. Juliant, S. *et al.* The α 1,6-fucosyltransferase gene (fut8) from the Sf9 lepidopteran insect cell line: insights into fut8 evolution. *PloS One* **9**, e110422 (2014).
20. O’Connell, K. P. *et al.* Production of a recombinant antibody fragment in whole insect larvae. *Mol. Biotechnol.* **36**, 44–51 (2007).
21. Düring, K., Hippe, S., Kreuzaler, F. & Schell, J. Synthesis and self-assembly of a functional monoclonal antibody in transgenic *Nicotiana tabacum*. *Plant Mol. Biol.* **15**, 281–293 (1990).
22. Houdebine, L.-M. Production of pharmaceutical proteins by transgenic animals. *Comp. Immunol. Microbiol. Infect. Dis.* **32**, 107–121 (2009).

23. Tang, B. *et al.* High level expression of a functional human/mouse chimeric anti-CD20 monoclonal antibody in milk of transgenic mice. *Transgenic Res.* **17**, 727–732 (2008).
24. Köhler, G. & Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**, 495–497 (1975).
25. Sgro, C. Side-effects of a monoclonal antibody, muromonab CD3/orthoclone OKT3: bibliographic review. *Toxicology* **105**, 23–29 (1995).
26. Morrison, S. L., Johnson, M. J., Herzenberg, L. A. & Oi, V. T. Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 6851–6855 (1984).
27. Nurden, P. *et al.* Delayed immunologic thrombocytopenia induced by abciximab. *Thromb. Haemost.* **92**, 820–828 (2004).
28. Lajus, S. *et al.* Thrombocytopenia after abciximab use results from different mechanisms. *Thromb. Haemost.* **103**, 651–661 (2010).
29. Butler, R. & Hubner, P. Acute severe thrombocytopenia after treatment with ReoPro (abciximab). *Heart* **83**, e5 (2000).
30. Giroux Portelance, S. Induction d’anticorps anti-idiotypiques contre les protéoglycanes de la matrice extracellulaire dans la réduction des lésions athérosclérotiques. (2013).
31. Soto, Y. *et al.* Antiatherosclerotic effect of an antibody that binds to extracellular matrix glycosaminoglycans. *Arterioscler. Thromb. Vasc. Biol.* **32**, 595–604 (2012).
32. Jones, T. D., Crompton, L. J., Carr, F. J. & Baker, M. P. Deimmunization of monoclonal antibodies. *Methods Mol. Biol. Clifton NJ* **525**, 405–423, xiv (2009).
33. Brüggemann, M. *et al.* A repertoire of monoclonal antibodies with human heavy chains from transgenic mice. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 6709–6713 (1989).
34. Frenzel, A., Schirrmann, T. & Hust, M. Phage display-derived human antibodies in clinical development and therapy. *mAbs* **8**, 1177–1194 (2016).

35. Laroche-Traineau, J., Clofent-Sanchez, G., Vezon, G. & Nurden, A. T. The ex vivo production of human monoclonal antibodies to glycoprotein IIb-IIIa complexes of blood platelets. *Hum. Antibodies Hybridomas* **5**, 165–177 (1994).
36. Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–1317 (1985).
37. McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* **348**, 552–554 (1990).
38. Lebani, K. *et al.* Isolation of serotype-specific antibodies against dengue virus non-structural protein 1 using phage display and application in a multiplexed serotyping assay. *PloS One* **12**, e0180669 (2017).
39. Tang, Y.-L. *et al.* Generation and Characterization of Antinonstructural Protein 1 Monoclonal Antibodies and Development of Diagnostics for Dengue Virus Serotype 2. *Am. J. Trop. Med. Hyg.* (2017). doi:10.4269/ajtmh.17-0003
40. Omar, N. *et al.* Generation and selection of naïve Fab library for parasitic antigen: Anti-BmSXP antibodies for lymphatic filariasis. *Biotechnol. Appl. Biochem.* (2017). doi:10.1002/bab.1591
41. Schenck, S. *et al.* Generation and Characterization of Anti-VGLUT Nanobodies Acting as Inhibitors of Transport. *Biochemistry (Mosc.)* **56**, 3962–3971 (2017).
42. Huang, Y. *et al.* An abundant dysfunctional apolipoprotein A1 in human atheroma. *Nat. Med.* **20**, 193–203 (2014).
43. Luque, M. C. A. *et al.* Phage Display Identification of CD100 in Human Atherosclerotic Plaque Macrophages and Foam Cells. *PLOS ONE* **8**, e75772 (2013).
44. Kumano-Kuramochi, M., Fujimura, T., Komba, S., Maeda-Yamamoto, M. & Machida, S. Screening, expression, and characterization of an anti-human oxidized low-density lipoprotein single-chain variable fragment. *J. Biosci. Bioeng.* **122**, 287–293 (2016).
45. Rondot, S., Koch, J., Breitling, F. & Dübel, S. A helper phage to improve single-chain antibody presentation in phage display. *Nat. Biotechnol.* **19**, nbt0101_75 (2001).

46. Welschof, M. *et al.* Amino acid sequence based PCR primers for amplification of rearranged human heavy and light chain immunoglobulin variable region genes. *J. Immunol. Methods* **179**, 203–214 (1995).
47. Barbié, V. & Lefranc, M. P. The human immunoglobulin kappa variable (IGKV) genes and joining (IGKJ) segments. *Exp. Clin. Immunogenet.* **15**, 171–183 (1998).
48. Pallarès, N., Frippiat, J. P., Giudicelli, V. & Lefranc, M. P. The human immunoglobulin lambda variable (IGLV) genes and joining (IGLJ) segments. *Exp. Clin. Immunogenet.* **15**, 8–18 (1998).
49. Pallarès, N., Lefebvre, S., Contet, V., Matsuda, F. & Lefranc, M. P. The human immunoglobulin heavy variable genes. *Exp. Clin. Immunogenet.* **16**, 36–60 (1999).
50. Barbas, C. F., Amberg, W., Simoncsits, A., Jones, T. M. & Lerner, R. A. Selection of human anti-hapten antibodies from semisynthetic libraries. *Gene* **137**, 57–62 (1993).
51. Barbas, C. F., Rosenblum, J. S. & Lerner, R. A. Direct selection of antibodies that coordinate metals from semisynthetic combinatorial libraries. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 6385–6389 (1993).
52. Hayashi, N. *et al.* Simultaneous mutagenesis of antibody CDR regions by overlap extension and PCR. *BioTechniques* **17**, 310, 312, 314–315 (1994).
53. Cobaugh, C. W., Almagro, J. C., Pogson, M., Iverson, B. & Georgiou, G. Synthetic antibody libraries focused towards peptide ligands. *J. Mol. Biol.* **378**, 622–633 (2008).
54. Söderlind, E., Vergeles, M. & Borrebaeck, C. A. Domain libraries: synthetic diversity for de novo design of antibody V-regions. *Gene* **160**, 269–272 (1995).
55. Soderlind, E., Carlsson, R., Borrebaeck, C. A. & Ohlin, M. The immune diversity in a test tube--non-immunised antibody libraries and functional variability in defined protein scaffolds. *Comb. Chem. High Throughput Screen.* **4**, 409–416 (2001).
56. Söderlind, E. *et al.* Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nat. Biotechnol.* **18**, 852–856 (2000).

57. Philibert, P. *et al.* A focused antibody library for selecting scFvs expressed at high levels in the cytoplasm. *BMC Biotechnol.* **7**, 81 (2007).
58. Vaughan, T. J. *et al.* Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nat. Biotechnol.* **14**, 309–314 (1996).
59. Little, M. *et al.* Generation of a large complex antibody library from multiple donors. *J. Immunol. Methods* **231**, 3–9 (1999).
60. Chapal, N. *et al.* Thyroid peroxidase autoantibodies obtained from random single chain FV libraries contain the same heavy/light chain combinations as occur in vivo. *Endocrinology* **142**, 4740–4750 (2001).
61. Jacobin, M.-J. *et al.* Human IgG Monoclonal Anti- $\alpha_{IIb}\beta_3$ -Binding Fragments Derived from Immunized Donors Using Phage Display. *J. Immunol.* **168**, 2035–2045 (2002).
62. Scholle, M. D., Collart, F. R. & Kay, B. K. In vivo biotinylated proteins as targets for phage-display selection experiments. *Protein Expr. Purif.* **37**, 243–252 (2004).
63. Keller, T. *et al.* Selection of scFv Antibody Fragments Binding to Human Blood versus Lymphatic Endothelial Surface Antigens by Direct Cell Phage Display. *PLoS One* **10**, e0127169 (2015).
64. Cao, J. *et al.* Phage display selection on whole cells yields a small peptide specific for HCV receptor human CD81. *Cell Res.* **13**, 473–479 (2003).
65. Ueberberg, S. *et al.* In vitro phage display in a rat beta cell line: a simple approach for the generation of a single-chain antibody targeting a novel beta cell-specific epitope. *Diabetologia* **53**, 1384–1394 (2010).
66. Cortese, I. *et al.* Identification of peptides specific for cerebrospinal fluid antibodies in multiple sclerosis by using phage libraries. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 11063–11067 (1996).
67. Casey, J. L., Coley, A. M., Parisi, K. & Foley, M. Peptide mimics selected from immune sera using phage display technology can replace native antigens in the diagnosis of Epstein–Barr virus infection. *Protein Eng. Des. Sel.* **22**, 85–91 (2009).

68. Ghoshal, S., Bondada, V., Saatman, K. E., Guttman, R. P. & Geddes, J. W. Phage display for identification of serum biomarkers of traumatic brain injury. *J. Neurosci. Methods* **272**, 33–37 (2016).
69. Govarts, C., Somers, K., Hupperts, R., Stinissen, P. & Somers, V. Exploring cDNA phage display for autoantibody profiling in the serum of multiple sclerosis patients: optimization of the selection procedure. *Ann. N. Y. Acad. Sci.* **1109**, 372–384 (2007).
70. Germaschewski, V. & Murray, K. Identification of polyclonal serum specificities with phage-display libraries. *J. Virol. Methods* **58**, 21–32 (1996).
71. Garlanda, C. & Dejana, E. Heterogeneity of endothelial cells. Specific markers. *Arterioscler. Thromb. Vasc. Biol.* **17**, 1193–1202 (1997).
72. Ruoslahti, E. & Rajotte, D. An address system in the vasculature of normal tissues and tumors. *Annu. Rev. Immunol.* **18**, 813–827 (2000).
73. Augustin, H. G., Kozian, D. H. & Johnson, R. C. Differentiation of endothelial cells: analysis of the constitutive and activated endothelial cell phenotypes. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **16**, 901–906 (1994).
74. Pasqualini, R. & Ruoslahti, E. Organ targeting in vivo using phage display peptide libraries. *Nature* **380**, 364–366 (1996).
75. Bussolati, B. *et al.* Targeting of human renal tumor-derived endothelial cells with peptides obtained by phage display. *J. Mol. Med. Berl. Ger.* **85**, 897–906 (2007).
76. Yang, P. *et al.* [Phage display random peptide library for screening the peptides that specifically bind to hepatocellular carcinoma cells]. *Nan Fang Yi Ke Da Xue Xue Bao* **29**, 1806–1808 (2009).
77. Chang, D.-K. *et al.* Antiangiogenic targeting liposomes increase therapeutic efficacy for solid tumors. *J. Biol. Chem.* **284**, 12905–12916 (2009).
78. Krag, D. N. *et al.* Selection of tumor-binding ligands in cancer patients with phage display libraries. *Cancer Res.* **66**, 7724–7733 (2006).

79. Arap, W. & Pasqualini, R. The human vascular mapping project. Selection and utilization of molecules for tumor endothelial targeting. *Haemostasis* **31 Suppl 1**, 30–31 (2001).
80. Pasqualini, R. *et al.* Targeting the interleukin-11 receptor α in metastatic prostate cancer: A first-in-man study. *Cancer* **121**, 2411–2421 (2015).
81. Arap, W. *et al.* Steps toward mapping the human vasculature by phage display. *Nat. Med.* **8**, 121–127 (2002).
82. Zurita, A. J. *et al.* Combinatorial screenings in patients: the interleukin-11 receptor alpha as a candidate target in the progression of human prostate cancer. *Cancer Res.* **64**, 435–439 (2004).
83. Qinghua, W. *et al.* A radiolabeled nonapeptide probe targeting PC-3 cells and bone metastases of prostate cancer in mice. *Contrast Media Mol. Imaging* **7**, 223–230 (2012).
84. Staquicini, F. I. *et al.* Vascular ligand-receptor mapping by direct combinatorial selection in cancer patients. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 18637–18642 (2011).
85. Kelly, K. A., Nahrendorf, M., Yu, A. M., Reynolds, F. & Weissleder, R. In vivo phage display selection yields atherosclerotic plaque targeted peptides for imaging. *Mol. Imaging Biol. MIB Off. Publ. Acad. Mol. Imaging* **8**, 201–207 (2006).
86. Deramchia, K. *et al.* In Vivo Phage Display to Identify New Human Antibody Fragments Homing to Atherosclerotic Endothelial and Subendothelial Tissues. *Am. J. Pathol.* **180**, 2576–2589 (2012).
87. Robert, R. *et al.* Identification of human scFvs targeting atherosclerotic lesions: selection by single round in vivo phage display. *J. Biol. Chem.* **281**, 40135–40143 (2006).
88. Reddy, S. T. *et al.* Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. *Nat. Biotechnol.* **28**, 965–969 (2010).
89. Ravn, U. *et al.* By-passing in vitro screening--next generation sequencing technologies applied to antibody display and in silico candidate selection. *Nucleic Acids Res.* **38**, e193 (2010).

90. Spiliotopoulos, A. *et al.* Sensitive recovery of recombinant antibody clones after their in silico identification within NGS datasets. *J. Immunol. Methods* **420**, 50–55 (2015).

Résultats

Chapitre 4: Résultats

Objectifs de la thèse

Le but principal du projet est de découvrir de nouveaux bio-marqueurs de l'athérosclérose par des techniques de phage display *in vivo* et de protéomique pour *in fine* développer des bio-traceurs afin d'être en mesure de visualiser la plaque d'athérome dite « vulnérable ». L'imagerie a non seulement pour but de diagnostiquer les sujets à risque mais également de suivre l'évolution de la plaque sous l'effet de substances médicamenteuses telles que la prostacycline, antiagrégant plaquettaire, que nous proposons comme principe actif pour nos agents théranostiques. La délivrance *in situ* d'une drogue par le bio-traceur sera assurée par des ligands de type anticorps humains qui possèdent une bonne affinité pour la cible.

L'objectif de ma thèse est, dans un premier temps, de cribler des anticorps humains reconnaissant des protéines surexprimées dans la plaque d'athérome *in vitro*. J'ai, pour cela, développé une technique de criblage à moyen débit par cytométrie de flux contre des protéines athéromateuses adsorbées sur des billes magnétiques. Les anticorps humains criblés par cytométrie de flux avaient été au préalable sélectionnés par phage display *in vivo* dans un modèle animal de l'athérosclérose, des lapins sous régime hypercholestérolémiant ayant subi deux chirurgies au niveau de l'aorte de façon à favoriser le développement de plaques complexes. Le phage display *in vivo* a été initié en 1996 par l'équipe de R.Pasqualini avec la sélection de peptides contre les vaisseaux sanguins rénaux ou hépatiques. Dans le projet de l'équipe dans laquelle j'ai effectué mes trois années de thèse, le phage display *in vivo* a été mis au point avec la sélection d'anticorps humains sous format scFv (single chain fragment variable). Nous avons utilisé pour cela une banque hyperdiversifiée construite par biologie moléculaire à partir de lymphocytes B de personnes saines et de patients atteints de diverses pathologies telles que les maladies auto-immunes ou cancers. Les anticorps humains ont été sélectionnés sur les protéines surexprimées dans la plaque d'athérome chez un lapin athéromateux par trois tours de sélection de manière à enrichir la sélection en phages spécifiques.

Dans un deuxième temps, l'objectif est d'étudier la possibilité de faire du criblage de ces anticorps humains *in silico* grâce aux dernières générations de séquençage (NGS). Le NGS apparaît aujourd'hui comme une nouvelle technique de criblage haut débit permettant de sélectionner et d'analyser des milliers de clones grâce au développement de puces avec plus d'un millier de puits. Le challenge reste cependant d'arriver à des séquences de scFv d'une taille supérieure à 1000pb comprenant deux V régions, ce qui n'a jamais été réalisé par le passé.

Dans un futur proche, les anticorps humains sélectionnés par ces deux méthodes seront couplés à des nano-objets pour immunocibler la plaque d'athérome. Cette étape consistera à les coupler à des agents de contraste super-paramagnétiques actuellement développés par l'équipe et des collaborateurs chimistes et pharmaciens pour visualiser la plaque d'athérome par imagerie par résonance magnétique (IRM).

Le projet de l'équipe « ciblage de la plaque d'athérome » au sein de laquelle j'ai effectué ces trois années de thèse a pour but ultime d'utiliser les avancées en ingénierie des anticorps et en imagerie non invasive pour le diagnostic et la thérapie, soit le théranostic de l'athérosclérose.

1^{ère} partie : Criblage *in vitro* d'anticorps

I Phage display *in vivo*

I.1 Introduction

A des stades différents de la progression de la pathologie, différents marqueurs biologiques peuvent être surexprimés dans la plaque d'athérome. A un stade précoce, certains marqueurs tels que les intégrines VCAM, ICAM sont exprimés après oxydation des LDL par les cellules endothéliales pour le recrutement cellulaire. A des stades plus tardifs, les marqueurs de l'apoptose, de la calcification sont exposés majoritairement ainsi que certaines enzymes telles que les métalloprotéinases ou les collagénases qui vont déstabiliser la paroi endothéliale et rendre la plaque susceptible de se rompre à tout moment. De ce fait, il est possible grâce à des modèles animaux développant des lésions athéroscléreuses, de cartographier *in vivo* l'état d'expression protéique depuis la formation de la strie lipidique jusqu'au stade où la plaque peut se rompre. Le phage display *in vivo* est parfaitement adapté à ce type d'usage. La sélection est réalisée sans connaissance préalable de la cible mais en fonction du niveau d'expression et de l'accessibilité du ligand. Pour rappeler le principe du phage display *in vivo*, il s'agit d'injecter une banque d'anticorps sous format single chain Fragment variable (scFv) exprimée à la surface de phages filamenteux dans la circulation sanguine du modèle étudié. Ainsi, les phages recombinants exprimant le scFv à leur surface vont se lier à leur ligand exprimé au sein de la lésion.

En comparaison avec une sélection *in vitro* contre un antigène ou une population cellulaire donnée, le ciblage *in vivo* donne un accès direct à des antigènes exposés dans leur conformation native non nécessairement présents *in vitro* mais également permet de découvrir de nouveaux bio-marqueurs d'une pathologie donnée. De plus, la sélection d'anticorps humains permet de pallier les problèmes d'immunogénicité que l'on peut rencontrer avec des anticorps murins ou humanisés. En dernier lieu, les anticorps sont d'une grande affinité pour une cible et peuvent être également déclinés sous plusieurs formats grâce à l'ingénierie des anticorps. Ainsi les trois avantages majeurs au phage display *in vivo* sont : (1) la sélection positive comme négative qui s'opère durant la circulation des phage-scFv, (2) le criblage *in vivo* contre des antigènes dans des conditions physiopathologiques et (3) l'identification de nouveaux bio-marqueurs ainsi que l'évaluation de leur niveau d'expression entre tissus sains et tissus athéroscléreux.

Les premières études *in vitro* de phage display faites dans le cadre de l'athérosclérose, en particulier sur le dysfonctionnement endothélial, se sont portées sur la sélection de peptides dirigés contre LOX-1, récepteur endothélial surexprimé dans l'athérogénèse, la P-selectine impliquée dans le rolling des leucocytes sur l'endothélium ou encore le récepteur CD100 exprimé sur les macrophages^{1,2,3}. Actuellement, essentiellement dû au problème de spécificité, les recherches s'orientent vers la sélection d'anticorps, molécules plus affines pour leur ligand. Citons notamment la sélection d'anticorps dirigés contre la plakoglobine, l'apoA1 modifiée ou encore les LDL oxydées^{4,5,6}. Bien que très peu d'études sur la sélection d'anticorps humains *in vivo* aient été rapportées dans la littérature, l'équipe au sein de laquelle j'ai effectué ma thèse a réussi à isoler deux anticorps contre la vitronectine, constituant de la matrice extracellulaire mais aussi la carbonic anhydrase II, impliquée dans la calcification de la plaque à un stade tardif^{7,8}. Ces deux études ont reposé sur l'injection dans la circulation sanguine de modèles animaux sous régime hypercholestérolémiant, de phage-scFv issus de banques semi-synthétiques. Après un temps défini de circulation, l'animal fût sacrifié, l'aorte extraite et les phage-scFv élués des zones d'intérêt. Pour chaque étude, la caractérisation des anticorps élués a été réalisée par criblage sur filtres ou par test HTRF (Homogeneous Time Resolved Fluorescence) respectivement^{7,8}.

I.2 Sélection *in vivo* de fragments d'anticorps

Basée sur les études précédentes, notre laboratoire a réalisé une sélection *in vivo* d'une banque immune par trois tours de biopanning en utilisant le modèle de lapin sous régime hypercholestérolémiant. Cette banque, en comparaison avec celles précédemment utilisées par l'équipe, présente l'avantage d'être issue d'immunoglobulines humaines de personnes saines et de patients atteints de diverses pathologies telles que le cancer, les maladies inflammatoires et les maladies auto-immunes. Les gènes d'immunoglobulines ont été extraits afin de construire par biologie moléculaire une banque de fragments scFv (single chain Fragment variable) incluant seulement les parties variables V_H et V_L d'une immunoglobuline. De manière à reproduire le processus d'hypermutation somatique, des modifications ont été apportées de manière aléatoire par l'utilisation de polymérases faible fidélité dites « mutases »⁹. Ainsi, après trois tours de sélection, les phage-scFv ont été criblés par une technique à moyen débit innovante, que j'ai mis au point au cours de ma thèse, la cytométrie de flux adaptée au criblage d'un pool de protéines extraites d'un tissu.

I.2.1 Le modèle animal utilisé : le lapin athéromateux

Le modèle lapin utilisé dans le cadre de la sélection *in vivo* de fragments d'anticorps est un lapin mis sous régime hypercholestérolémiant pendant 6-8 mois. Afin de faciliter l'apparition de plaques et de reproduire la composition d'une plaque humaine, le lapin a subi deux interventions chirurgicales à 1-2 mois d'intervalle¹⁰. La première intervention consiste en une désendothélialisation de la paroi artérielle par gonflement d'un ballon de manière répétée au niveau de la partie basse de l'aorte thoracique descendante et de l'aorte abdominale. La deuxième intervention permet d'aller aplatiser la plaque à l'intérieur de l'artère grâce à un ballon extensible en latex. Il s'agit ici d'une angioplastie au niveau de la partie basse de l'aorte thoracique descendante.

I.2.2 La sélection *in vivo* réalisée

La sélection a été réalisée par trois tours de biopanning chez le modèle lapin athéromateux. A chaque tour de sélection, les phages ont été élués de façon fractionnée dans différentes sections et fractions de l'aorte, amplifiés puis réinjectés à un nouvel animal athéromateux (**Figure 47**).

Brièvement, 10¹² phages ont été injectés dans la circulation sanguine du modèle lapin athéromateux. Après 1h de circulation, l'animal fût sacrifié et l'aorte, organe d'intérêt où les plaques se forment de manière préférentielle, a été disséquée. Le but étant ici de récupérer les phages fixés sur des cibles présentes en surface mais également internalisées, trois fractions ont été éluées : F1 correspondant à la fraction de phages fixés sur la surface endothéliale ; F2 : correspondant à la fraction de phages intra-plaque ; F3 : correspondant à la fraction de phages internalisés dans les cellules résidentes de la plaque. Deux chirurgies ayant été réalisées au niveau de la partie basse de l'aorte thoracique descendante et abdominale pour favoriser la formation des plaques, les phages ont été élués de quatre sections différentes. E1.1 prend en compte la crosse aortique. La section E1.2 correspond à la partie haute de l'aorte thoracique descendante. E2.1 correspond à la partie basse de l'aorte descendante jusqu'aux bifurcations rénales. La section E2.2 correspond à la section entre les bifurcations rénales et les bifurcations iliaques. La partie E1 est donc une partie non abrasée de l'aorte alors que la partie E2 correspond à la partie abrasée de l'aorte. Les phages ainsi élués ont été à chaque tour amplifiés dans le système bactérien de manière à conserver l'ADN des clones d'intérêt. Lors de ma thèse, j'ai utilisé ce matériel génétique et testé la réactivité de plus de mille clones uniques par cytométrie de flux, moyen de criblage à moyen débit.

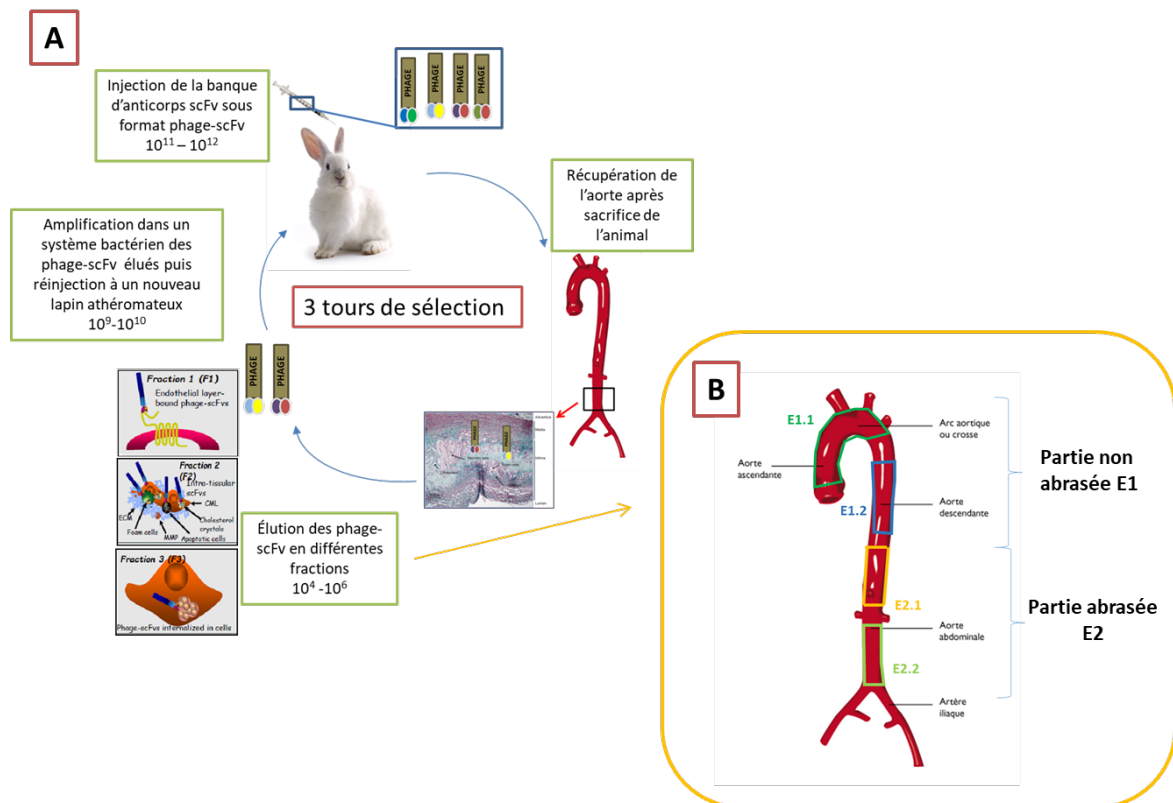


Figure 47 Sélection par phage display *in vivo* d'une banque d'anticorps humains. (A) Représentation schématique du phage display *in vivo*. Trois cycles de biopanning ont été effectués chez le lapin NZW athérosclérotique. Lors d'un tour de sélection, les phage-scFv sont injectés dans la circulation sanguine du lapin. Après une heure de circulation, l'animal est sacrifié et l'aorte est récupérée. Les phage-scFv ont ensuite été élués de différentes fractions : F1: phage-scFv récupérés sur la surface endothéliale, F2: phage-scFv récupérés au niveau intratissulaire et F3: phage-scFv récupérés au niveau intracellulaire. (B) Les phage-scFv ont également été élués de plusieurs sections, notamment E1.1, E1.2, E2.1, E2.2.

I.2.3 Production de phages recombinants

Afin de tester la réactivité par cytométrie de flux des clones sélectionnés *in vivo*, les clones bactériens ont été mis en présence de phages auxiliaires afin de produire des phages recombinants par la méthode de phage display.

I.2.4 Test de réactivité

I.2.4.i Cytométrie de flux

Cette méthode permet de faire défiler des complexes en suspension dans un milieu liquide devant un laser afin de les caractériser. La cytométrie de flux permet donc de différencier les complexes marqués des complexes non marqués par analyse du signal optique émis par le complexe lors du passage devant le faisceau lumineux du laser. Ainsi la molécule est caractérisée par sa taille (Forward Scatter), sa granulométrie (Side Scatter) et son intensité relative de fluorescence. Les signaux lumineux sont séparés par des filtres optiques puis collectés par des photomultiplicateurs, amplifiés, numérisés et traités pour être stockés dans l'ordinateur (**Figure 48**). D'un point de vue historique, la cytométrie en flux tient ses origines des années 30, utilisée en premier lieu par Andrew Moldovan pour le comptage de cellules¹¹.

Les phage-scFv produits ont été criblés par la technique de cytométrie en flux, mise au point au cours de ma thèse. En effet, l'un des problèmes majeurs inhérents à la technologie du phage display est que lors de la production de phages recombinants, une majorité d'entre eux, environ 95% n'expriment pas à leur surface le fragment d'anticorps. Ces phages recombinants sans scFv exprimé apportent un bruit de fond important lors du criblage qu'il a été nécessaire de réduire. Pour cela, plusieurs conditions de lavage et d'incubation ont été testées (représentées en **Figure 49**). Après la sélection de la condition déterminée comme optimale, les phage-scFv ont été criblés par cytométrie de flux.

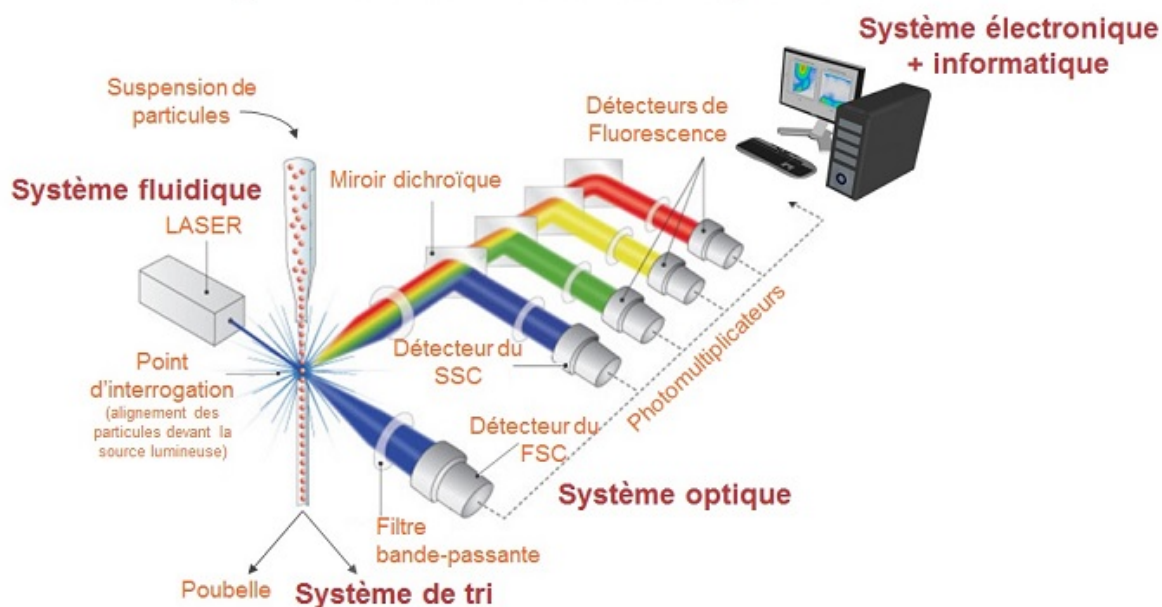


Figure 48 Principe de fonctionnement de la cytométrie de flux. Brièvement, les molécules sont séparées suivant leur taille (SSC), leur granulométrie (FSC) et leur fluorescence. Les signaux lumineux sont séparés par des filtres, collectés par des photomultiplicateurs puis numérisés par un système informatique. FSC : Forward Scatter ; SSC : Side Scatter (Extrait de <http://www.univ-reims.fr/site/plateaux-techniques/plateau-technique-de-cytometrie-en-flux>)

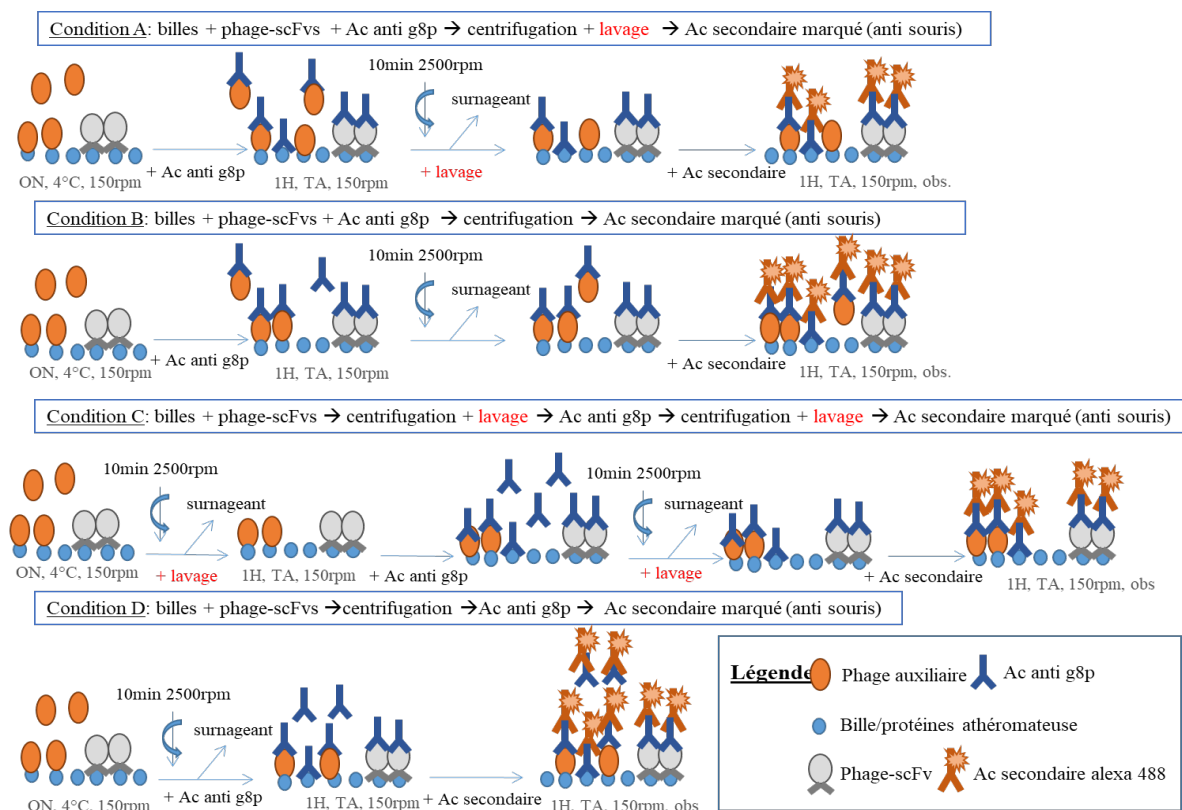


Figure 49 Schématisation des conditions (A à D) testées en cytométrie de flux pour la mise au point du criblage de phage-scFv. Plusieurs conditions ont été testées. Ces dernières diffèrent par la position des lavages et des centrifugations avant ou après l’ajout des anticorps qui permettent la révélation. Juste deux des conditions A et D sont présentées dans l’article.

1.2.4.ii Immunohistochimie

Tous les clones ayant été sélectionnés par cytométrie de flux ont été testés en immunohistochimie pour leur capacité à reconnaître leurs cibles sur des tissus athéromateux isolés du modèle lapin.

1.3 Résultats et discussion

L’ensemble des résultats sont décrits de manière détaillée dans l’article intitulé « An innovative flow cytometry method for the screening of *in vivo* selected scFv-phages issuing from a fully human antibody library “ soumis dans la revue *Scientific Reports* et dont je suis l’auteur principal.

Lors de cette étude, j’ai tenté de sélectionner des anticorps humains (qui sont ainsi moins susceptibles d’engendrer une réponse immune lors d’une injection ultérieure chez l’homme) et qui présentent également par rapport aux peptides une meilleure affinité pour une cible donnée. Dans le but de sélectionner ces anticorps dans un contexte physiopathologique, nous avons développé une méthode de phage display *in vivo* chez un lapin de la pathologie étudiée, un lapin athéromateux. Ainsi, après trois tours de sélection, environ 3.8×10^7 phage-scFv ont été récupérés et individualisés en masterblock (**Tableau 4**).

Tableau 4. Tableau récapitulatif de l'enrichissement des phage-scFvs au cours des différents tours de sélection (P1 – P3) dans les différentes fractions. Le nombre de phage-scFv a été normalisé par rapport à la quantité de phages injectée, soit 10^{12} phage-scFv (afin de comparer les cycles itératifs de biopanning *in vivo*), et calculé par mg de tissus.

	P1	P2	P3
E1F1	4,36E+06	1,03E+05	1,96E+06
E2F1	7,44E+05	3,28E+05	9,02E+05
E1F2	1,38E+06	7,62E+04	7,20E+06
E2F2	2,83E+06	6,98E+05	2,52E+07
E1F3	1,52E+04	4,37E+03	1,61E+06
E2F3	1,19E+05	N/A	1,13E+06

De manière à pouvoir ensuite les caractériser et confirmer leur bio-réactivité, j'ai implémenté une technique de criblage à moyen débit, la cytométrie de flux. Du fait que ces anticorps aient été sélectionnés *in vivo*, les cibles sont inconnues et présentes, lors du criblage, en quantité infime parmi d'autres protéines dans des extraits athéromateux. L'utilisation d'un criblage par ELISA étant donc impossible, le développement d'une technique faisant appel à une quantité de protéines faible a été nécessaire. Après mise au point du procédé de cytométrie de flux, 200 anticorps, parmi plus de 2000 clones criblés, ont pu être sélectionnés puis séquencés par la technologie classique de Sanger et testés par immunohistochimie afin de valider le protocole de criblage. Soixante pourcent des clones sélectionnés par cytométrie de flux et séquencés par Sanger ont montré une réactivité *ex vivo* sur coupes d'aorte de lapin athéromateux.

En parallèle du criblage de phage-scFv issus de la sélection *in vivo* chez un lapin athéromateux, une étude protéomique de la plaque d'athérome chez ce modèle animal d'athérosclérose a permis de mettre en exergue une multitude de cibles potentielles pouvant être les ligands des phages sélectionnés. Grâce à l'extraction des protéines athéromateuses, j'ai pu conduire une étude sur la composition de la plaque d'athérome développée chez le lapin et identifier les mécanismes reliés à ces protéines.

Lors de l'analyse, les ratios des abondances normalisées des protéines athéromateuses provenant de lapins sous régime hypercholestérolémiant sur les abondances normalisées des protéines issues de lapins sains ont été calculés et importés dans Pathway Studio. Deux tests ont été réalisés : (1) le test GSEA (Gene set enrichment analysis) prenant en compte les ratios de toutes les protéines et (2) le test FISHER ne prenant en compte uniquement que le Top100 des protéines les plus représentées.

Cette étude a permis dans un premier temps de valider notre modèle d'étude en vérifiant l'appartenance des protéines et leur implication dans des processus cellulaires connus dans la database pour leur lien avec l'athérosclérose (**Figure 50**). Comme nous avons pu l'observer, une majorité des processus cellulaires était en lien avec la réponse immunitaire impliquant les macrophages, les neutrophiles et les lymphocytes, démontrée dans la littérature comme prépondérante dans l'athérogénèse. Une partie des processus était également rattachée au métabolisme des lipides qui est le principal acteur de la formation de la plaque d'athérome. De plus, les lipides ingérés par les macrophages sont à l'origine de la formation des cellules spumeuses qui lorsqu'elles rentrent en apoptose font que le cœur lipidique et nécrotique se rompe et entraîne des maladies cardio-vasculaires graves. Nous avons pu également voir la présence de processus liés au stress oxydatif, à la morphogénèse vasculaire ou encore la coagulation du sang.

Dans un deuxième temps, nous avons focalisé notre attention sur le métabolisme des lipides. Ainsi, à l'intersection de plusieurs sous-processus cellulaires se trouvent des protéines connues pour leur rôle dans l'athérogénèse (**Figure 50A**). En effet, comme répertoriée dans la littérature scientifique, l'Apolipoprotéine A1 (APOA1), composant majeur des HDL, permet l'efflux de cholestérol inhibant ainsi l'accumulation des cellules spumeuses au sein de la plaque. De plus, en interaction avec la paroxonase 3 (PON3), ces deux protéines ont un effet anti oxydant important. Le complexe possède la capacité à inhiber la modification des LDL, réduisant l'effet athérogène de ces lipoprotéines ¹². L'Apolipoprotéine E (APOE) est impliquée dans le métabolisme du cholestérol. Il s'agit d'un transporteur des lipides. Il a été montré que la déficience en APOE conduit à l'hyperlipoprotéïnémie et une dysfonction du catabolisme du cholestérol ¹³.

Si on s'intéresse ensuite aux protéines les plus exprimées (protéines en rouge dans la **Figure 50B**) dans le sous ensemble « foam cell formation », la cathepsine D (CTSD) qui colocalise avec le lysosome des macrophages, participe à la modification de l'ApoB par oxydation. Les LDL ainsi oxydées sont reconnues puis phagocytées par les macrophages ¹⁴. La carboxylestérase 1 (CES1) est impliquée dans le métabolisme du cholestérol et des acides gras. Elle est associée aux macrophages. Au cours de l'athérogénèse, cette protéine serait hydrolysée ce qui induit une formation de cellules spumeuses ¹⁵. L'acétyl-CoA-acétyltransférase (ACAT1) est une enzyme qui agit dans l'accumulation du cholestérol intracellulaire et l'assemblage des lipoprotéines ¹⁶. Le modificateur d'ubiquitine 1 (UFM1), également localisé au niveau des macrophages, semble inhiber la formation des cellules spumeuses par la voie LXR α ¹⁷. Le CD36, comme vu précédemment, est un récepteur scavenger impliqué dans l'accumulation des LDL ox dans les macrophages, participant donc à la formation des cellules spumeuses ¹⁸. La protéine associée aux acides gras (FABP4), exprimée par les macrophages et les adipocytes est décrite comme ayant un rôle important dans le métabolisme, le transport et l'ingestion des acides gras ¹⁹. La protéine reliée aux récepteurs de LDL (LRP1) est un récepteur aux LDL, lié à l'homéostasie des lipides et la clairance des cellules apoptotiques ²⁰. La cavéoline 1 (CAV1) a été montrée pour être impliquée dans l'inflammation. En effet, une déficience de cette protéine chez des souris conduit à une réduction de l'infiltration des LDL, promeut la production de NO et réduit l'expression de molécules d'adhésion influençant le recrutement des cellules immunitaires ²¹.

Pour finir, si on s'attache plus particulièrement au processus cellulaire conduisant à la formation des cellules spumeuses, certaines protéines peuvent être mises en exergue en tant que biomarqueurs de la plaque d'athérome. L'intégrine β 2 (ITGB2), également nommée CD18, participe à l'adhésion des cellules et à la signalisation cellulaire. Cette protéine est exprimée à la surface des leucocytes et interagit avec la molécule d'adhésion ICAM-1 exprimée à la surface de l'endothélium. Ce processus permet la diapédèse des leucocytes qui correspond à la migration de ces cellules à travers l'endothélium et donc leur entrée au site inflammatoire ²². La prothrombine (F2), quant à elle, agit dans la formation du thrombus à un stade tardif de la plaque d'athérome ²³.

Ainsi, parmi tous ces biomarqueurs pouvant être les cibles des anticorps sélectionnés *in vivo* puis *in vitro*, le challenge futur sera de pouvoir découvrir la cible exacte de chacun grâce au reformatage des anticorps sous format scFv-Fc dans le système BaculobMab ou dans le système de cellules mammifères (HEK293). Ainsi, cela nous permettra d'avoir une meilleure appréhension de l'athérogénèse et donner accès à des anticorps pouvant être couplés à des nano-objets utilisables en imagerie moléculaire de la plaque d'athérome.

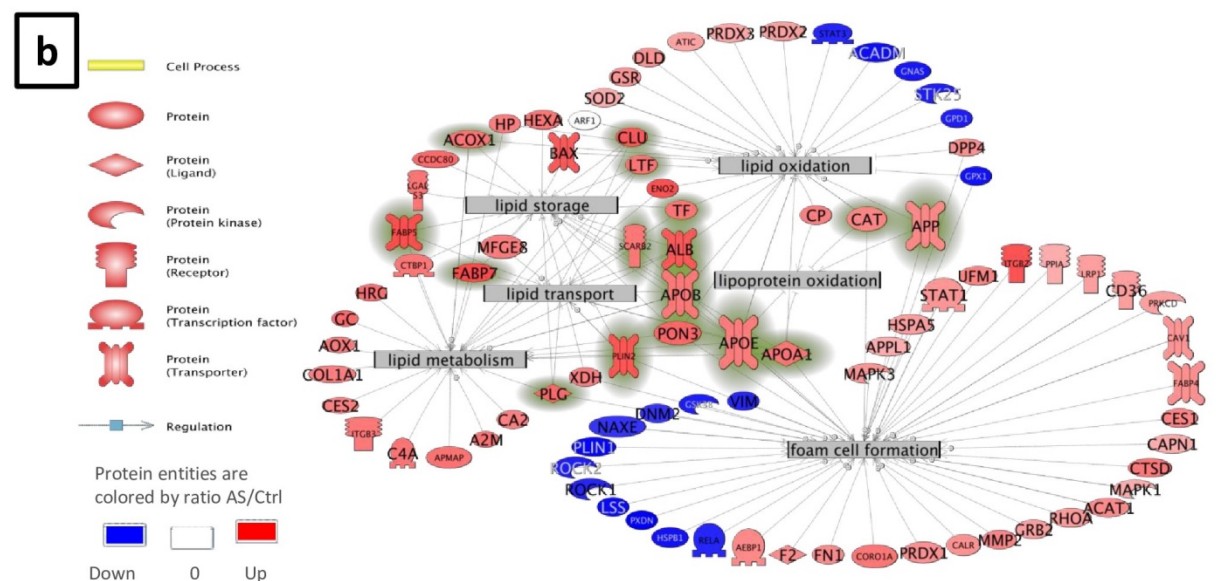
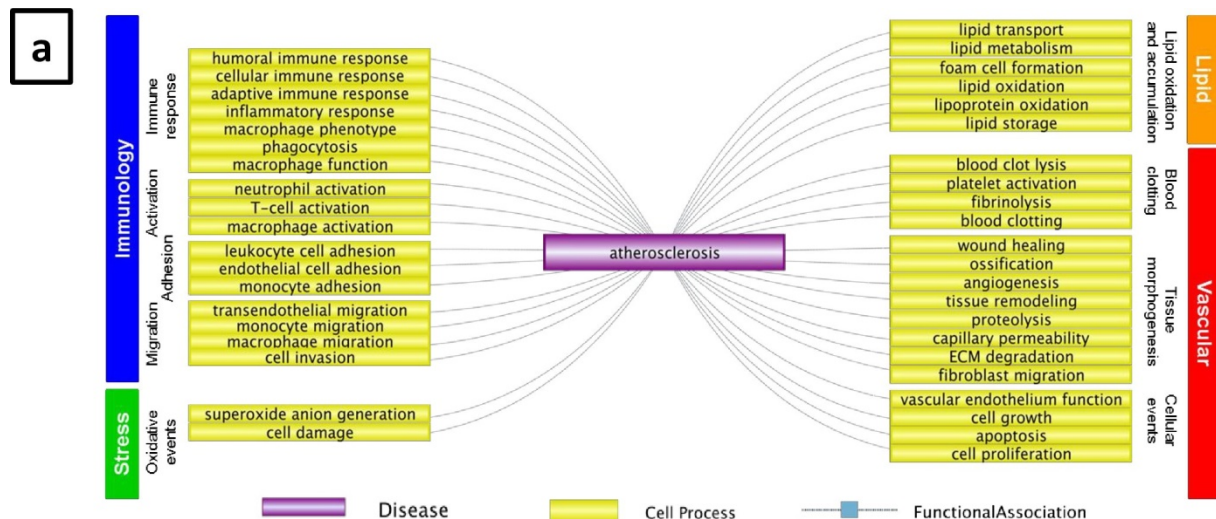


Figure 50 Sous-réseaux de processus cellulaires enrichis dans des échantillons d'aorte de lapins NZW athéroscléreux comparés à des lapins témoins NZW. a: Analyse des processus cellulaires liés aux protéines surexprimées dans la plaque athéromateuse de lapins NZW hypercholestérolémiques. Les principaux processus cellulaires signalés dans Pathway Studio® à la fois fonctionnellement associés à l'athérosclérose et régulés par de multiples protéines avec une abondance différentielle dans les extraits protéiques aortiques de lapins NZW athérosclérotiques et sains sont rapportés. Les processus cellulaires sont regroupés en 4 catégories: événements immunitaires, vasculaires, liés au stress et aux lipides. Tous les liens sont soutenus par des résultats biologiques de la littérature scientifique. b: Sous-réseaux enrichis liés aux processus des cellules lipidiques. Les entités protéiques colorées en rouge et en bleu sont régulées respectivement à la hausse et à la baisse dans les extraits protéiques d'aorte d'athérosclérose par rapport aux lapins NZW sains. Les protéines surlignées en vert corégulent dans 3 ou plus des 6 processus cellulaires. Tous les liens sont soutenus par des résultats biologiques de la littérature scientifique. La variation de la taille de police du nom de l'entité n'a aucune signification biologique.

Criblage de phage-scFv *in vitro* par une
technique innovante, la cytométrie de
flux à moyen débit

(Article)

An innovative flow cytometry method for the screening of *in vivo* selected scFv-phages issuing from a fully human antibody library

Audrey HEMADOU¹, Jeanny LAROCHE-TRAINEAU^{†1}, Ségolène ANTOINE^{†1}, Philippe MONDON², Alexandre FONTAYNE², Yannick LE PRIOL³, Stéphane CLAVEROL⁴, Stéphane SANCHEZ¹, Martine CERUTTI⁵, Florence OTTONES¹, Gisèle CLOFENT-SANCHEZ^{†1} and Marie-Josée JACOBIN-VALAT^{†*1}

[†]Equivalent position of authors

¹CRMSB, UMR5536 CNRS, INSB, Bordeaux, 33076, France

²LFB Biotechnologies, department of biotherapeutic engineering Lille, 59000, France

³Elsevier Masson SAS, Elsevier RD solutions, Issy les Moulineaux, 92130, France

⁴CGFB, Proteome pole, Bordeaux, 33076, France

⁵UPS 3044, CNRS, Saint-Christol-Lès-Alès, France

*Correspondence

Marie-Josée JACOBIN-VALAT

marie-josee.jacobin-valat@rmsb.u-bordeaux2.fr

Abstract

Atherosclerosis is a chronic, progressive inflammatory disease that may develop into vulnerable lesions leading to thrombosis. This pathology is characterized by the deposition of lipids within the arterial wall and infiltration of immune cells leading to amplification of inflammation. Nowadays there is a rising interest to assess directly the molecular and cellular components that underlie the clinical condition of stroke and myocardial infarction. Single chain fragment variable (scFv)-phages issuing from a human combinatorial library were selected on the lesions induced in a rabbit model of atherosclerosis after three rounds of *in vivo* phage display. We further implemented a high-throughput flow cytometry method on rabbit protein extracts to individually test one thousand of scFv-phages. Two hundred and nine clones were retrieved on the basis of their specificity for atherosclerotic proteins. Immunohistochemistry assays confirmed the robustness of the designed cytometry protocol. Sequencing of candidates demonstrated their high diversity in VH and VL germline usage. The large number of candidates and their diversity open the way in the discovery of new biomarkers. Here, we successfully showed the capacity of coupling *in vivo* phage display and high-throughput cytometry strategies to give new insights in *in vivo* targetable up-regulated biomarkers in atherosclerosis.

1 Introduction

Atherosclerosis is the main cause of death in the western world with 19 million deaths per year¹. This chronic and systemic disease affects arteries with large and medium caliber². The pathology is characterized by the build-up of lipid-rich plaques within the arterial wall which can evolve into stable or vulnerable atheroma plaques. Stable atheroma tends to be asymptomatic unless expanding atherosclerotic lesions cause severe narrowing of the lumen, one of the consequences being ischemia. Vulnerable atheroma has more dramatic impacts due to the presence of a large lipid core covered with a thin fibrous cap at high risk of rupture and thrombi formation. These events precipitate the clinical conditions of stroke and myocardial infarction^{3,4}. Nowadays, high-grade

internal carotid artery stenosis (>70% luminal narrowing) is awaited before proposing endarterectomy to the patients⁵. However, the degree of luminal stenosis is not always effective in identifying high-risk patients and less than 50% occlusion in the arterial bed can lead to dramatic events. Only targeted imaging modalities able to characterize atherosclerosis at the molecular level could provide a valuable mean to diagnose vulnerability. Even though progress in understanding the atherogenic molecular basis has been made, opening up horizons for several promising novel targets such as $\alpha\text{IIb}\beta\text{3}$ ⁶ and carbonic anhydrase II (CAII)⁷, there is still a crucial need to develop clinical imaging modalities with new ligands able to target efficiently the atherogenic components⁸.

In this study, in order to assess the cellular and molecular components underlying the risk of plaque rupture in pathophysiological conditions, we proceed to an *in vivo* phage display biopanning in an animal model of atherosclerosis. Initially, *in vivo* phage display selections had been performed with peptide libraries^{9,10,11}. We recently implemented this biotechnological approach with human scFv (single-chain fragment variable) libraries because human antibodies provide better specificity and lower immunogenicity when injected in humans. Human antibody fragments were successfully selected and shown to target new biomarkers of the pathology^{7,12,13}. However, these studies used semi-synthetic human scFv libraries that typically yield antibody fragments of lower affinity. This contrasts the situation in immune libraries where the immune system proceeds to affinity maturation for specific antigens. In order to meet affinity and specificity requirements and thus a better efficacy of targeting, we switched to a fully human scFv antibody (HuAb) library with high and original complexity made of a wide naive HuAb population combined with a naturally “oriented” population of HuAb against several pathologies (autoimmune diseases, different types of cancer and cardiovascular diseases). This human library (MG-Umab) was further hyper-diversified by random mutagenesis (patent N° WO0242311, WO2007137616)^{14,15} in order to generate 3.4 billion clones with 94% of them encoding full-length variable genes. Here, we described the successful selection of scFv-phages obtained from this fully human library by three rounds of *in vivo* phage display biopanning in atherosclerotic New Zealand White (NZW) rabbits combined with high-throughput flow cytometry screening. This method is highly suitable to test thousands of *in vivo* selected scFv-phages. Individually selected phage-antibodies whose specificity was confirmed by immunohistochemistry (IHC) will be of interest in the discovery of novel targets for a better understanding of the pathogenesis of atherosclerosis and the development of accurate diagnosis tools.

2 Results

The successive selection steps to identify scFv-phage candidates able to target cellular and molecular atherogenic components are summarized in Figure 1. The whole strategy is based on the coupling of two methodologies: *in vivo* phage display and high-throughput flow cytometry. Briefly, scFv-phages issued from a fully human library were exposed to atherogenic cellular and molecular components of rabbit aorta by an *in vivo* approach. Individual clones isolated during this *in vivo* selection were then screened by FACS analysis against balloon injured aortas of hypercholesterolemic NZW rabbit (BIAHR) proteins. After quantification of scFv-phage by ELISA, western blotting was performed for further control of scFv-pIII expression. Selected hits were validated by IHC for bioreactivity assays.

2.1 *In vivo* phage display selection of human scFvs in a complex animal model of atherosclerosis

ScFv-phages isolated from the lesions (from the aortic arch to the thoracic aorta and from the reanl to the illiac bifurcations) constitute the round 1-selected scFv-phage library (Figure2a and b). Iterative cycles of *in vivo* selection and amplification (additional round 2 and round 3) in subsequent

NZW rabbits were performed to enrich for more specific candidates (Figure 2a). Furthermore, in order not to restrict the analysis only on sites directly in contact with blood, scFv-phages emigrating within the intima were recovered. Homing scFv-phages were thus isolated not only from the injured endothelium (F1 fraction) but also from the underlying lesional tissue (F2 fraction) and within the cells from the intima (F3 fraction) (Figure 2c). The number of scFv-phages recovered (cfu) was normalized for an input of 10^{12} cfu (in order to compare the iterative cycles of *in vivo* biopanning) and calculated per mg of tissue. The number of recovered scFv-phages from F2 and F3 fractions increased after the third round of selection compared to the first round (Supplementary Figure S1). The implemented technology has thus led to the *in vivo* selection of scFv-phages able to reach subjacent lesional tissue layers and internalized within cells invading the intima.

2.2 High-throughput biological screening

One of the drawbacks of *in vivo* selections is the screening strategy of the selected clones. Indeed, the antigen is not identified and that limits the use of classic ELISA assays where purified antigens are usually coated on plates at a quite high concentration (10 $\mu\text{g}/\text{mL}$). Our solution was to implement a high-throughput flow cytometry plaque-assay using soluble tissue extracts coupled to magnetic beads supposed to be more sensitive compared to ELISA with less biological material needed. Indeed, only 200 ng of proteins (see Methods section) are used per well to screen individual clones by flow cytometry, meaning 5 fold less than for an ELISA assay. Moreover, high-throughput screening should allow testing thousands of clones obtained from *in vivo* biopanning.

2.2.1 Pathway Studio analysis of proteins differentially expressed in atherosclerotic NZW rabbits

It was firstly necessary to assess the diversity and over-expression of proteins in the BIAHR extracts. Because of the lack of data in the rabbit database, the protein extraction composition was compared to the human database. The ratios of proteins between BIAHR and Healthy (H) extracts were calculated from the normal abundance and then imported and studied by an international software, Pathway studio[®].

2.2.1.1 *The proteomic profile of balloon-injured aortas in hypercholesterolemic NZW rabbits reveals atheromatous characteristics*

The study of the data issued from the proteomic analysis was performed following two approaches, one considering the entire data set using the GSEA-based (SNEA) test¹⁶, the other focusing on the top 100 up-regulated proteins analysed by Fisher's based (FSNE) assay¹⁷. In a first step, ratios of the whole protein profiles and the top 100 upregulated proteins were compared to the pathway database to find regulated cell processes. The analysis showed that the overexpressed proteins in BIAHR compared to H aortas are implicated in a large diversity of cell processes that have been related to either immunological-, vascular-, stress- or lipid-related events (Figure 3a). Most of these events were described to contribute to the natural history of atherosclerosis^{2,18} (Supplementary Table S1 and Figure 3a). In a second step, the question was: *what cell processes are linked to atherosclerosis according to the Pathway Studio[®] database?* The analysis demonstrated that almost half of the top cell processes found by GSEA and FSNE tests were also found to be functionally associated with atherosclerosis in the Pathway Studio[®] knowledge base of millions of biological findings from peer-reviewed literature (Supplementary Table S1). Thus, first of all, these data gave us confidence in the composition of our BIAHR extracts. Given the involvement of the immune response

in the context of atherogenesis, it was not surprising that immunological cell processes were over-represented (81/178) and many of them related to the activation, migration or adhesion of immune cells. Interestingly, some of enriched cell processes were linked to lipids (18/178), stress (2/178) or events implying vessels (42/178) such as blood clotting, tissue morphogenesis and more general cell processes related to cell growth, cell proliferation and apoptosis (Figure 3a).

2.2.1.2 Zoom on proteins involved in lipid-related cell processes

We choose to focus on lipid-related process because of the high involvement of lipids in atherogenesis. The protein subnetwork regulating lipid oxidization (Figure 3b) was found significantly enriched in both Fisher's ($p=5.7 \times 10^{-3}$) and GSEA enrichment test ($p=1.7 \times 10^{-9}$). The one regulating foam cell formation was found significantly enriched only in GSEA enrichment test (GSEA $p=7.1 \times 10^{-3}$). It was shown here that some proteins are implicated and regulated in more than one lipid-related subnetwork. Some of them are well known to be increased in atherosclerosis such as apolipoprotein E (ratio=10.82), ApoB (ratio=14.53) and APOA1 (ratio=17.33) as well as paraoxonase 3 (PON3) (ratio=45.74)^{19,20} (Chistiakov et al). The most upregulated proteins in BIAHR compared to H involved in foam cell formation were the integrin $\beta 2$ (ratio =147), Perilipin 2 (PLIN2) (ratio=132) (Pourcet et al), Plasminogen (PLG) (ratio=37.7) (Kremen et al), Coronin, actin-binding protein 1A (CORO1A) (ratio=16.8), Cathepsin D (CTSD) (ratio=14.1), Xanthine dehydrogenase, (XDH) (ratio=8.6), Carboxylesterase1 (CES1) (ratio=6.5), Ubiquitin fold modifier 1 (UFM1) (ratio=6.3), Catalase (CAT) (ratio=5.8), Matrix metalloprotease 2 (MMP2) (ratio=5.65) (Table 1). All these results highlighted over-expressed proteins which may be considered as new biomarkers of atherogenesis, potentially targeted by human antibodies issued from our *in vivo* selection. Other proteins are down-regulated in the foam cell formation sub-network such as Rho associated protein kinase (ROCK2) (ratio=0.7), Glycogen synthase kinase 3 beta (GSK3B) (ratio=0.7), Heat shock 27kDa protein 1 (HSPB1) (ratio=0.5), Glutathione peroxidase 1 (GPX1) (ratio=0.5), Lanosterol synthase (LSS) (ratio=0.3) and Perilipin 1 (PLIN1) (ratio=0.2).

2.2.2 Evaluation of the coupling of BIAHR proteins onto the beads

For an efficient screening of scFv-phages by flow cytometry, it is important to know the rate of coupling of the protein extracts to the magnetic beads. Thus, the three batches of coated protein beads were analyzed by SDS-PAGE electrophoresis. In electrophoresis conditions, only the uncoupled fractions of protein-beads preparations were able to migrate through the gel matrix, the beads coated with BIAHR proteins being stacked on top of the gel. In Figure S2 (given in supplementary data), the signal obtained with the uncoupled fraction is identical between the three batches and in the same range as this obtained with 8 to 10 μg of the whole extract loaded on wells. The coupling rate was estimated by Image J analysis at approximately 80% (Supplementary Figure S2). We designed a first flow cytometry assay (Figure 4a) to evaluate, after coating, the integrity and the folding of the proteins by using AP2 and CAII murine antibodies. These antibodies were chosen for their well-known capacity to bind proteins present in the atherosclerotic extract with different ratio BIAHR on H proteins (13.69 for integrin $\beta 3$ and 8.37 for carbonic anhydrase II) (see the ratio of proteins in Supplementary Table S2). The results obtained assessed the quality of the coupling and accredited the whole flow cytometry strategy. Indeed, as shown in Figure 4a, the increase of the

fluorescence intensities of AP2 and CAII over the secondary antibodies confirmed their binding on their respective targets (α IIb β 3 integrin and the carbonic anhydrase II enzyme).

2.3 Determination of accurate screening conditions

The phage display technology has the inconvenience to amplify a large amount of recombinant phages that do not display the scFv fragment because they are more able to infect bacteria with a high efficacy (95 % of phages without scFv expression have been estimated by using M13KO7 helper phage). The hyperphage used in this study, in contrast to M13KO7 helper phage, was designed to force the *E. coli* cell to use the pIII portion of the antibody fusion protein for packaging. However, a significant amount of phages still not display scFv fragments. These phages may bring about non-specific binding during the screening process which is necessary to reduce. It was thus mandatory to set up the conditions for FACS analysis to minimize the non-specific binding of hyperphage particles on coated beads. For that, we used different concentrations of hyperphage particles from 10^8 to 10^{10} pfu/mL in accordance to the range of scFv-phage production classically obtained in masterblock cultures.

As shown in Figure 4b in condition D, the fluorescence intensity increases when using hyperphages concentrations from 10^9 to 10^{10} pfu/mL, highlighting a non-specific binding linked to the concentration itself. This non-specific binding could be explained by the lack of centrifugation and washing after the addition of anti-pVIII antibody (Table 2). We thus introduced these steps in condition C. The problem remained with the concentration of 10^{10} pfu/mL giving a high background signal. We tested another strategy by removing centrifugation and washing steps after overnight hyperphages/proteins incubation, before anti-pVIII antibody incubation (conditions A and B). Surprisingly, this notably reduced the background signal. An hypothesis could be that the excess of anti-pVIII antibody may be washed out by hyperphage particles not withdrawn between overnight hyperphages/proteins incubation and addition of anti-pVIII antibody. This wash-out may thus reduce the recognition of the hyperphage without scFv bound non-specifically on atheromatous proteins coated on beads (Supplementary Figure S3). We finally chose condition A in which only one centrifugation and one washing were performed to remove the residual anti-pVIII antibody and anti-pVIII antibody bound to hyperphage before addition of the secondary antibody. To conclude, in condition A, the variation is minimal whatever the concentration of hyperphage particles used, showing a minimal non-specific binding. This is an important point considering that at the time of the FACS experiment the rate of production of each clone could not be determined. Indeed, to be in line with a high-throughput strategy, it is impossible to quantify by ELISA each clone of a 96 deep-well culture plate. ELISA experiment will only be conducted for flow cytometry selected hits for further validation steps.

2.4 Screening of phage antibodies by flow cytometry on BIAHR proteins

Eight hundred sixty four clones from the third round of biopanning were produced from F1, F2 and F3 fractions in 96 deep-well masterblocks (PL19 to PL27) (supplementary Table S3). They were tested, by flow cytometry following previously implemented condition A, for their reactivity against atheroma protein extracts. Overall, 209 clones (24 %) reacted positively with atherosclerotic proteins by binding about 2 fold or more over the threshold signal produced by hyperphage. This threshold was arbitrary fixed by the geomean obtained on all positive events defined as a P2 gate. Figure 5 shows typical examples of fluorescent intensities of eight positive and negative clones from PL20 measured by their geomean value calculated on the P2 gate. We can see that a good correlation was obtained between this theoretical definition of positivity by geomean comparison and the switch of

the fluorescence towards higher values for positive recombinant phages. The peaks were quite different from each other, considering the differences in terms of representativity of the target on the beads as well as productivity of each scFv-phage clone. Here, the panel of fluorescence histograms summarizes the diversity of reactivity of clones, some of them being categorized as highly positive and others with a positivity just above the threshold. It is also illustrated a clone which was considered as negative. For the sake of clarity, only the results of one of the three batches are shown, knowing that the results within the three batches are always highly reproducible in each experiment. It is important to underline that this high-throughput flow cytometry strategy is a pre-screening approach, which needs further validations of the candidates by (1) sequencing the scFv fragments to control the presence of entire VH and VL genes (2.5 section) and (2) evaluating the reactivity of scFv-phage clones within the context of atherosclerosis burden by immunohistochemistry (IHC) assays on atherosclerotic NZW rabbit sections (2.7 section). Before IHC experiments, the quantification of the scFv-phages production rate in masterblocks and the accurate display of scFv-pIII on top of the hyperphage determined by ELISA and Western blotting respectively, are necessary (2.6 section). Some of the clones which were non-reactive in flow cytometry were also introduced in the assays as controls.

2.5 ScFv antibody sequencing and bioinformatic analysis

In order to control and estimate the diversity of the selected clones that bind atherosclerotic proteins, their VH and VL genes were sequenced and analyzed by IMGT/V-QUEST (Supplementary Methods S1). This allowed us to evaluate the integrity of the whole scFv with in frame CDR3 junctions (Supplementary Table S3). Among the 209 clones selected by flow cytometry, 68% of them have a complete sequence with 42% without stop codons. Analyses of the variable germlines and CDR3 junctions showed a high diversity among the clones, excluding a potential biased clone enrichment (Supplementary Table S3). In conclusion, the scFv sequence analysis underlies a potential diversity of paratopes and unbiased clone abundance in the round 3 of *in vivo* phage selection, highlighting a potential diversity in targeted biomarkers.

2.6 Quantification of phage production and expression of g3p-scFv after hyperphage superinfection

We recall that, in spite of the use of hyperphage, during expression of the scFv-pIII, the supernatant contains both wild-type and recombinant phages. So, in order to evaluate the display of pIII-scFv on top of the recombinant phages, selected clones were analyzed by western blotting after ELISA quantification of the whole production (wild-type and recombinant phages) (Supplementary Methods S2). The concentration of eight clones individually produced was estimated by ELISA in regards to a range of hyperphage particles (10^8 to 10^{11} pfu/mL). For these clones, the concentrations varied between $5 \cdot 10^9$ and 10^{10} pfu/mL. These concentrations higher than the loading threshold in western blotting (around 10^8 phages) allowed their analysis to detect the scFv-pIII expression. Figure S4 (given in Supplementary Information) shows that all the clones expressed the wild-type pIII and the pIII-scFv protein. In contrast to what is described in the literature, the band representing the scFv-pIII is lower than what is observed for the wild-type pIII. This is explained by the use of a plasmid, to build the scFv-library, that contains a truncated pIII gene (pIII_t) (patent WO2007137616). We can notice that there is a great variability of expression level of the scFv-pIII protein between the different clones. Indeed, D3PL20, C9PL20, C10PL20 and B6PL20 presented a stronger expression of the scFv-pIII than E3PL20, C3PL20, B7PL20 and A2PL20. This is especially true for E3PL20 for which a

tiny band at the scFv-pIII level is observed. We can also notice that the chosen negative clone A2PL20 showed a moderate expression of the scFv-pIII protein. All along the study, for a given clone, we observed from batch to batch differences in production level (Supplementary Figure S2). In spite of unequal expression profiles, we decided to pursue with all the clones to determine their bioreactivity on atherosclerotic sections keeping in mind that the reactivity may be impacted the expression level besides unequal representativeness of the targets.

2.7 Evaluation of the selected scFv-phages reactivity on atherosclerotic NZW rabbit tissue sections

The IHC experiments conducted on eight representative clones with atherosclerotic NZW rabbit tissue sections are shown in Figure 6. Hyperphages were used as a control at a concentration of $5 \cdot 10^9$ phages/mL, determined by ELISA for the eight clones tested (Supplementary Methods S3). B6PL20, C3PL20, C9PL20, C10PL20 and E3PL20 clones (Fig 6c, e, f, g, i) bound preferentially the sub-endothelial layer whereas B7PL20 and D3PL20 (Fig 6d, h) showed a broader reactivity within the intima. We noticed that B7PL20 and D3PL20 reacted with intimal proteins with the same intensity despite a very different expression level of scFv-pIII, underlying the fact that scFv expression is not strikingly correlated to the level of binding. Thus, we show here that even poorly displayed scFv can give rise to good reactivity in IHC experiments, depending on either the representativeness of the target or the affinity of the scFv. Whatever, the IHC experiments corroborate the results obtained with the high-throughput flow cytometry assay. Considering all of our experiments (PL19 to PL27), IHC analyses revealed that 60% of clones, selected by flow cytometry, were able to recognize over-expressed proteins within sections of atheroma plaque. The forty percent of clones that did not recognize their target in IHC were withdrawn from further studies. This fail in antigen recognition may result from false positive results in flow cytometry but also from the modification of epitopes occurring after treatment of tissues with paraformaldehyde fixation, and/or high temperatures treatment of the tissue sections. On the contrary, we have experienced some clones initially found negative during the flow cytometry screening step showing in IHC assays tiny reactivity on atheromatous sections. We were initially aware of this pitfall due to limitations in the extraction process, essentially caused by the loss of some proteins and the degradation of some others. Flow cytometry experiment using protein extracts is not a conformational screening strategy, meaning that some scFv could be missed by this approach. However, this strategy has the benefit to offer a higher throughput capacity compared to IHC. Altogether, IHC certifies the reactivity of these pre-screened clones on sections of atheroma and acknowledges the robustness of the flow cytometry screening.

3 Discussion

Acute cardiovascular events, due to rupture or erosion of an atherosclerotic plaque, represent the major cause of morbidity and mortality in patients. Growing evidence suggests that plaque vulnerability is directly linked to its molecular and cellular composition, which now places new demands on molecular imaging. It is expected from this rapidly evolving field an accurate assessment of molecular, cellular and compositional components of atherosclerotic lesions to efficiently diagnose high-risk patients. This imaging modality requires highly sensitive and specific probes made of an affinity ligand for targeting and a signal detection compound.

Magnetic resonance imaging (MRI), a non-invasive modality, has a great potential for clinical application considering multiple advantages such as its high resolution and the absence of ionization²¹. Thanks to the knowledge on some biomarkers overexpressed in atherogenesis, researchers developed functionalized contrast agents dedicated to magnetic resonance imaging to target $\alpha\text{v}\beta\text{3}$ integrin overexpressed in angiogenesis²², fibrin involved in thrombosis²³, annexin V expressed on apoptotic macrophages²⁴, oxLDL²⁵ and scavenger receptor from foam cells²⁶ or P-selectin on activated platelets and endothelial cells²⁷. These MRI studies were mainly conducted with nanoparticles coupled to peptides or commercial and home-made murine and rat antibodies.

Because of two major inconveniences, which are the immunogenicity of rodent Abs and the lower affinity of peptides for their targets, the development of human antibodies to gain affinity and decrease immunogenicity is the aim of this study. In the context of atherosclerosis, designing human antibodies as affinity ligands for targeting is very new, except sparse studies of Fab fragments raised against well-defined antigens such as oxLDL²⁸. Here another novelty which we consider important to emphasize is the selection of these human antibodies by *in vivo* phage display for accessing a large panel of targets in pathophysiological conditions. Because the environment impacts the folding of proteins and their expression, *in vitro* methods often result in the selection of peptides or antibodies no more able to recognize their target when transferred *in vivo* for pre-clinical and clinical studies. Thus, to get around this issue, we implemented *in vivo* phage display selections that may result in an array of recombinant scFv-phage clones with the potential to recognize various and unknown targets. By extending the selection over three rounds, we aimed to force the enrichment of specific antibodies. This strategy led to a strong increase in scFv-phage titers after the third round.

Because the selection of HuAbs has been made *in vivo*, a crucial step remains the screening for individual scFv-phage clones against unknown proteins. A big challenge was the set-up of a high-throughput screening with low quantities of biological material meanwhile making available a high surface density of different target antigens. The fact that the antigen is unknown limits the use of classic ELISA assays where purified antigens are usually coated on plates at quite high concentrations. Here, we demonstrate that it is highly feasible to screen around thousand of scFv-phage clones by the design of a suitable high-throughput method using flow cytometry and atheromatous proteins coated on beads at low concentration. Flow cytometry binding experiments attested to enrichment of functional clones and allowed the selection of a lot of scFv-phage hits, which presented substantiated binding for the three batches of protein-coated beads.

Out of two hundred and nine clones selected by flow cytometry, we retrieved 142 clones with complete and in frame VH and VL germline genes after sequencing of the whole scFv fragment. Moreover, analysis by IMGT/V-QUEST has highlighted a high diversity of germline genes and CDR3 junctions, pointing out a potential high diversity in targeting biomarkers. IHC experiments with selected scFv-phage clones have been included (1) to prove the robustness of flow cytometry screening with the reactivity of 60% of selected clones confirmed *ex vivo* on NZW rabbit atheroma sections and (2) to provide visual localisation of these pre-screened clones within atheroma.

Moreover, in this work, proteomic analyses have highlighted hundreds of proteins overexpressed in the context of atherogenesis, validating the atheromatous model used here to screen human antibodies and defining potential new biomarkers. The Pathway Studio® analysis has identified a lot of proteins upregulated in the atherosclerosis context and well-known in the literature such as apolipoproteins E, B and A1, integrins, scavenger receptors. On the other side, the analysis has highlighted the presence of proteins, not well assigned in the literature to atherogenesis which will be interesting to test against selected antibodies. The next challenge will be to identify among the

panel of atheromatous proteins in the extract, the targeted molecules, making use of selected scFv produced as scFv-Fc fragments at large-scale.

These high-scale produced human scFv-Fc antibodies could thus help to survey the molecular and cellular components of atheroma plaques, increasing our understanding of different molecular pathways (as seen by proteomics analyses) leading to the development of atheroma.

In future studies, the human scFv-Fc antibodies will serve for *in vivo* experiments on atherosclerotic ApoE^{-/-} mice as suitable moieties after grafting on nanoparticles for (i) a more accurate diagnosis of atherosclerosis dedicated to MRI imaging and (ii) the development of therapeutic strategies. It is to be emphasized that human antibodies which have been selected in a relevant animal model of atherosclerosis, would allow direct transfer from preclinical to clinical studies. The human antibodies that will be used in pre-clinical studies will be those able to target the corresponding human proteins on human atheroma sections.

The strategy of selection by *in vivo* phage display coupled to high-throughput screening by highly sensitive flow cytometry which has the capability to recognize proteins underrepresented in the extract could be applied to many other studies in cancer and auto-immune diseases for identification of target/ligand pairs.

4 Methods

4.1 Complex plaque formation in the NZW rabbit animal model

All animal experiments were performed in accordance with the Guide for the Care and Use of Laboratory Animals (NIH Publication No. 85-23, revised 1996) and were approved by the ethic committee of Bordeaux (CEEA50). The NZW rabbit model was fed with a high cholesterol diet as previously described¹³ and two surgeries were performed, two months apart. The first consisted in a desendothelialization from the thoracic until the abdominal aorta and the second in an angioplasty carried out with a balloon through a femoral arteriotomy to the descendent thoracic aorta under radiosopic guidance.

4.2 *In vivo* phage display selection in the atherosclerotic NZW rabbit model

4.2.1 ScFv-phage library preparation

The fully human recombinant fragment scFv antibody library described in patent WO2007137616 was used for *in vivo* phage display selection. For phage rescue, the phage library was infected with M13-KO7 helper phage (Invitrogen, France), followed by precipitation of phages with PEG/NaCl. After centrifugation at 11,000 g for 45 min at 4 °C, the pellet containing scFv-phage particles was resuspended and filtered in a final volume of 500 µL sterile cold PBS.

4.2.2 *In vivo* phage-display

Three rounds of biopanning were performed by phage display in atherosclerotic NZW rabbits. The full scheme of the procedure is given in Figure 2. The first selection was carried out using a continuous flow (150 µL/min) of 2.4×10^{12} colony-forming units (cfu) of scFv-phages injected into the NZW rabbit marginal ear vein for 30 min. After 1h circulation, the animal was sacrificed and perfused via the heart with 120 mL of PBS to ensure phage clearance from the blood. The aorta was recovered from the aortic arch to the iliac bifurcations and divided into E1 (aortic arch and thoracic area) and E2

(abraded renal and abdominal areas) sections as shown in Figure 2. Rounds 2 and 3 were conducted following the same procedure, excepted that the quantity of scFv-phages was lowered to 4.8×10^{11} and 3.9×10^{11} cfu respectively.

To access not only the scFv-phages binding to the endothelial cell surface but also those migrating into the intima and those internalized into cells invading the lesions, we respectively recovered F1, F2 and F3 fractions from aortic tissue (Figure 2).

Details concerning the recovery of the different fractions are given in the supplementary data section (Supplementary Methods S4).

ScFv-phage fractions F1, F2 and F3 were separately rescued by infecting XL1-blue *E. coli* (Stratagene, France) (OD=0.5) and the cells were pelleted and plated on ampicillin (100 µg/mL) and glucose (2 % (w/v)) containing 2xTY agar (2TYGA) and incubated 16 h at 30 °C. Spreading on 145 mm Petri dishes has allowed constituting the glycerol stock and limiting dilutions on 80 mm dishes has allowed evaluating the enrichment of selection. After the third round of biopanning, plated clones were individually picked into deep well 96 masterblocks filled with 2TYGA selective medium (Greiner Bio one, France) for FACS screening and further analyses.

4.3 Protein extraction and mass spectrometry study

4.3.1 Protein extraction

The protein extraction was performed as previously described⁷. Briefly, BIAHR and H proteins were solubilized with a commercially available T-PER lysis buffer (Thermo Fisher Scientific, France) complemented with a protease inhibitors cocktail (Thermo Fischer scientific, France). Solubilization was performed using a Polytron TP-20 Homogenizer 8 (Kinematica, Lucerne, Switzerland). After two centrifugations, at 13,000 g for 45 min at 4 °C to discard insoluble material from the supernatant, the protein concentration of every soluble extract was determined using a Bradford assay kit according to the manufacturer's instructions (Thermo Fisher Scientific, France).

4.3.2 Mass spectrometry of protein extract

The diversity of proteins in the aorta extracts was assessed by mass spectrometry at the Centre de génomique fonctionnelle of Bordeaux.

4.3.2.1 nLC-MS/MS analysis and database search

Ten µg of proteins were loaded in Laemlli buffer onto SDS-PAGE under reducing condition using 10% gels. Separation was stopped once proteins have entered resolving gel. Colloidal blue stained bands were cut out from the gel and treated as described elsewhere²⁹. Briefly, gel pieces were destained and digested with trypsin overnight at 37 °C. Peptides were extracted and injected on a Ultimate 3000 nanoLC system (Dionex, Amsterdam, The Netherlands) coupled to a Electrospray Q-Exactive mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Mass spectrometry data were acquired in a data-dependent mode alternating a MS scan survey on m/z 300 – 2000 with HCD (High Collision Dissociation) MS/MS spectra on top 15 ions. Data were searched by SEQUEST through Proteome Discoverer 1.4 (Thermo Fisher Scientific Inc, France) against a subset of the 2014.05 version of UniProt database restricted to *Homo sapiens* Reference Proteome Set (68,358 entries). Tolerances were set to 10 ppm and 0.02 Da for precursors and fragments respectively. Only b- and y-ions were considered for mass calculation. Oxidation of methionines and deamidation of asparagines and glutamines were considered as variable modifications and carbamidomethylation of cysteines as

fixed modification. Two missed trypsin cleavages were allowed. Peptides were validated at a False Positive Rate of 1 % using Percolator⁵. Label-free quantitative data analysis was performed using Progenesis LC-MS 4.1 software (Nonlinear Dynamics Ltd, U.K). Importantly, normalization was based on median ratio, and only non-conflicting features and unique peptides were considered for calculation. Quantitative data were considered for proteins quantified by a minimum of 2 peptides.

4.3.2.2 Functional analysis

Protein profiles of the aortic arch and thoracic aorta from healthy and balloon-injured aortas of hypercholesterolemic NZW rabbit were compared. Ratios of normalized protein abundance between BIAHR and H extracts were calculated and imported in Pathway Studio[®] 11.3 for functional analysis and biological meaning. GSEA-based (SNEA) and Fisher's based (FSNE) cell-process subnetwork enrichment analysis tests were performed using default settings on the entire ratio dataset and the list of top 100 upregulated proteins respectively. Only identified cell processes also known to be functionally associated with the atherosclerosis disease entity in the Pathway Studio knowledgebase were retained for further categorisation and interpretation. The sentences of all links supported by less than three references in enriched subnetworks were closely reviewed, and only highly confident links were kept.

4.4 ScFv-phage preparation

Individual recombinant XL1-blue colonies from the third biopanning were grown overnight at 30 °C and shaken at 280 rpm (Newbrunswick, Edison, USA), in 96-deep well plates (masterblock) containing 500 µL of 2TYGA medium supplemented with 10 µg/mL of tetracyclin (Tet). Twenty five microliters of this previous culture were used to inoculate 500 µL of 2TYGAT medium for the scFv-phage production. After 3 hours of bacteria growth, phage production was induced by adding 25 µL of 2TY medium containing 3×10^8 hyperphage (Progen, Germany). After 1h of infection at 37°C, bacteria were then pelleted and supernatants discarded. Infected bacteria were subsequently resuspended in 500 µL of 2TYA medium supplemented with 40 µg/mL kanamycin (Kan) for a last growth overnight at 26 °C, 260 rpm. Bacteria were finally spun down at 10,000 g, 10 minutes and supernatants were immediately used for FACS analyses.

4.5 Phage quantification by ELISA

Wells of flat bottom corning costar 96-well EIA/RIA plates (Fischer Scientific, France) were coated overnight at 4 °C with mouse anti-M13 (specific to pVIII protein) monoclonal antibody (GE healthcare, U.K) diluted at 5 µg/mL in carbonate buffer (15 mM carbonate, 35 mM bicarbonate/ pH 9.6). Wells were then blocked with 5 % MPBS (milk PBS). Dilutions of scFv-phage supernatants (1:10 and 1:100) and a range of hyperphage (Progen, Germany) from $5 \cdot 10^8$ to 10^{11} cfu/mL were added to wells in duplicates and incubated for two hours at RT. After washing, HRP-conjugated anti-pVIII protein monoclonal antibody (GE healthcare), diluted 1:5,000 in 2 % MPBS, was added for 1 hour at RT. Each step was followed by extensive washing in PBS Tween-20 0.5 % buffer (TPBS). The final washing was carried out with PSB alone and color was developed with OPD system for ELISA (Sigma Aldrich, France). Color absorbance was immediately read at 405 nm in a plate reader (Chameleon, Thermofisher, France).

4.6 FACS analyses

4.6.1 Coupling of proteins to magnetic beads

Fifty micrograms of BIAHR proteins were covalently coupled to 300 nm carboxyl-adenbeads according to the manufacturer instructions (Ademtech, France). Three batches of coated protein beads per sample were used for reproducibility. The coupling rate of BIAHR proteins on carboxyl-adenbeads was analyzed on SDS-PAGE under reducing condition using 4-15 % gels (miniprotean TGX gel, Biorad, France). The rate of uncoated fractions which migrated in the gel was determined by using the Image J software.

4.6.2 ScFv-phage screening by flow cytometry

Specificity of scFv-phages to BIAHR proteins was determined by flow cytometry in 96-well plates (Becton Dickinson, France). Forty microliters of BIAHR proteins coated on beads were added at 5 µg/mL to 100 µL of phage particles (range of hyperphage from 10^8 to 10^{10} cfu/mL) or individual scFv-phage and were incubated overnight at 4 °C. Different experimental conditions (A, B, C and D) were first tested to evaluate the rate of non-specific binding of hyperphage particles on coated protein-beads. These conditions, summarized in Table 2, differ in the number of centrifugations (10 min at 8,000 g) and washes performed before or after addition of a murine anti-pVIII primary antibody (Abcam, France) at 1 µg/mL for one hour at RT, 150 rpm. The Alexa 488 labeled anti-mouse antibody (Life technologies, France) was used at 3.75 µg/mL for final phage detection. To evaluate the coupling efficiency and quality of proteins on beads, murine AP-2 antibody (courtesy of Dr Alan Nurden) (anti- α IIb β 3 antibody) and rabbit anti-carbonic anhydrase II (CAII) antibody (Abcam, France) were incubated in the same conditions as above with the difference that secondary Alexa 488 labeled antibodies were used at 5 µg/mL (goat anti-mouse) or 3 µg/mL (goat anti-rabbit) for the detection of AP2 or CAII respectively. Samples were analyzed on a BD FACS Canto I device (Becton Dickinson, France) using software Diva. All washes and dilutions were performed in PBS. Fluorescence geomean was statistically calculated on the basis of 10,000 fluorescent events.

References

1. Naghavi, M. *et al.* From Vulnerable Plaque to Vulnerable Patient. *Circulation* **108**, 1664–1672 (2003).
2. Lusis, A. J. Atherosclerosis. *Nature* **407**, 233–241 (2000).
3. Chan, J. M. S., Monaco, C., Wylezinska-Arridge, M., Tremoleda, J. L. & Gibbs, R. G. J. Imaging of the Vulnerable Carotid Plaque: Biological Targeting of Inflammation in Atherosclerosis using Iron Oxide Particles and MRI. *Eur. J. Vasc. Endovasc. Surg.* **47**, 462–469 (2014).

4. Tarin, C. *et al.* Targeted gold-coated iron oxide nanoparticles for CD163 detection in atherosclerosis by MRI. *Sci. Rep.* **5**, (2015).
5. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S. & MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **4**, 923–925 (2007).
6. Jacobin, M.-J. *et al.* Human IgG Monoclonal Anti- $\alpha_{IIb}\beta_3$ -Binding Fragments Derived from Immunized Donors Using Phage Display. *J. Immunol.* **168**, 2035–2045 (2002).
7. Deramchia, K. *et al.* In Vivo Phage Display to Identify New Human Antibody Fragments Homing to Atherosclerotic Endothelial and Subendothelial Tissues. *Am. J. Pathol.* **180**, 2576–2589 (2012).
8. Sanz, J. & Fayad, Z. A. Imaging of atherosclerotic cardiovascular disease. *Nature* **451**, 953–957 (2008).
9. Shiu, D. *et al.* Exploration of Peptide Inhibitors of Human Squalene Synthase through Molecular Modeling and Phage Display Technique. *Appl. Biochem. Biotechnol.* **178**, 312–323 (2016).
10. Hofmeister, L. H. *et al.* Phage-display-guided nanocarrier targeting to atheroprone vasculature. *ACS Nano* **9**, 4435–4446 (2015).
11. Luque, M. C. A. *et al.* Phage Display Identification of CD100 in Human Atherosclerotic Plaque Macrophages and Foam Cells. *PLOS ONE* **8**, e75772 (2013).
12. Robert, R. *et al.* Identification of human scFvs targeting atherosclerotic lesions: selection by single round in vivo phage display. *J. Biol. Chem.* **281**, 40135–40143 (2006).
13. Deramchia, K. *et al.* By-Passing Large Screening Experiments Using Sequencing as a Tool to Identify scFv Fragments Targeting Atherosclerotic Lesions in a Novel In Vivo Phage Display Selection. *Int. J. Mol. Sci.* **13**, 6902–6923 (2012).
14. Mondon, P. *et al.* Method for generation of human hyperdiversified antibody fragment library. *Biotechnol. J.* **2**, 76–82 (2007).
15. Mondon, P., Bouayadi, K., Kharrat, A. & Millegren. *Bibliothèques d'anticorps hautement diversifiées.* (2006).

16. Irizarry, R. A., Wang, C., Zhou, Y. & Speed, T. P. Gene Set Enrichment Analysis Made Simple. *Stat. Methods Med. Res.* **18**, 565–575 (2009).
17. Jung, S.-H. Stratified Fisher's Exact Test and its Sample Size Calculation. *Biom. J. Biom. Z.* **56**, (2014).
18. Libby, P. Inflammation in Atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **32**, 2045–2051 (2012).
19. Barter, P. J. *et al.* Antiinflammatory Properties of HDL. *Circ. Res.* **95**, 764–772 (2004).
20. Emini Veseli, B. *et al.* Animal models of atherosclerosis. *Eur. J. Pharmacol.*
doi:10.1016/j.ejphar.2017.05.010
21. Choudhury, R. P., Fuster, V., Badimon, J. J., Fisher, E. A. & Fayad, Z. A. MRI and Characterization of Atherosclerotic Plaque. *Arterioscler. Thromb. Vasc. Biol.* **22**, 1065–1074 (2002).
22. Winter, P. M. *et al.* Molecular Imaging of Angiogenesis in Early-Stage Atherosclerosis With $\alpha_v\beta_3$ -Integrin-Targeted Nanoparticles. *Circulation* **108**, 2270–2274 (2003).
23. Sirol, M. *et al.* Chronic Thrombus Detection With In Vivo Magnetic Resonance Imaging and a Fibrin-Targeted Contrast Agent. *Circulation* **112**, 1594–1600 (2005).
24. Kietselaer, B. L. J. H. *et al.* Noninvasive Detection of Plaque Instability with Use of Radiolabeled Annexin A5 in Patients with Carotid-Artery Atherosclerosis. *N. Engl. J. Med.* **350**, 1472–1473 (2004).
25. Briley-Saebo, K. C. *et al.* Targeted Molecular Probes for Imaging Atherosclerotic Lesions With Magnetic Resonance Using Antibodies That Recognize Oxidation-Specific Epitopes. *Circulation* **117**, 3206–3215 (2008).
26. Amirbekian, V. *et al.* Detecting and assessing macrophages in vivo to evaluate atherosclerosis noninvasively using molecular MRI. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 961–966 (2007).
27. Jacobin-Valat, M.-J. *et al.* MRI of inducible P-selectin expression in human activated platelets involved in the early stages of atherosclerosis. *NMR Biomed.* **24**, 413–424 (2011).

28. Shaw, P. X. *et al.* Human-Derived Anti-Oxidized LDL Autoantibody Blocks Uptake of Oxidized LDL by Macrophages and Localizes to Atherosclerotic Lesions In Vivo. *Arterioscler. Thromb. Vasc. Biol.* **21**, 1333–1339 (2001).
29. Fessart, D. *et al.* Proteomic remodeling of proteasome in right heart failure. *J. Mol. Cell. Cardiol.* **66**, 41–52 (2014).

Acknowledgements:

We would like to thank Chloé DELELA who contributed to this work. This study was funded by (1) Grant ANR-13-BSV5-0018 French National Research Agency Program named ATHERANOS, (2) a public grant from the French National Research Agency within the context of the Investments for the Future Program, referenced ANR-10-LABX-57 and named TRAIL.

Author contributions:

Audrey HEMADOU conceived and designed the experiments. She wrote the paper.

Jeanny LAROCHE-TRAINEAU conceived and designed the experiments.

Ségolène ANTOINE conceived and designed the experiments.

Philippe MONDON provided the MGHumab phage library and the technical support to use it.

Alexandre FONTAYNE contributed by his exchange to the content and the writing of the manuscript.

Yannick LE PRIOL contributed by his exchange to the content and the writing of the manuscript, concerning the proteomic analysis on Pathway Studio Software.

Stéphane CLAVEROL contributed by his exchange to the content and the writing of the manuscript on the proteomic analysis and provided proteomic data.

Stéphane SANCHEZ conceived and designed the experiments.

Martine CERUTTI contributed by his exchange to the content and the writing of the manuscript.

Florence OTTONES contributed by his exchange to the content and the writing of the manuscript.

Gisèle CLOFENT-SANCHEZ contributed by his exchange to the content and the writing of the manuscript and supervised the project.

Marie-Josée JACOBIN-VALAT conceived and designed the experiments and contributed to the writing of the manuscript and supervised the project.

All the authors have read and approved the final manuscript.

Additional information:

Competing Interests : The authors declare no competing financial interests.

Figure legends:

Figure 1. Representative scheme of the developing process of antibodies able to recognize atheroma plaques. (A) Human scFv-phages were injected in atherosclerotic NZW rabbits for *in vivo* recognition of cellular and molecular atherogenic components in the aorta. After *in vivo* biopanning, scFv-phages were retrieved from rabbit aorta. (B) FACS was used to individually screen scFv-phages issuing from *in vivo* phage-display against atherosclerotic proteins. (C) The clones were sequenced to evaluate the integrity of the two V-Regions. (D) Quantitative ELISA was used to quantify individual scFv-phage clones produced in 96-well culture plates. (E) scFv-pIII protein expression was controlled by Western Blotting. (F) The bioreactivity of human scFv-phage candidates was validated on atherosclerotic sections by immunohistochemistry.

Figure 2. Schematic representation of the *in vivo* phage display procedure. a: Three rounds of biopanning were performed in atherosclerotic NZW rabbit. ScFv-phages were eluted from different fractions of the extracted aorta. b: Fractions issued from two sections of the aorta, E1 and E2 sections. E1 section included the aortic arch and the thoracic area and the E2 section corresponded to renal and abdominal areas which were abraded by balloon injury. c: F1: recovered scFv-phages from endothelial fraction, F2: recovered scFv-phages from intra-tissular fraction and F3: recovered scFv-phages from intra-cellular fraction.

Figure 3. Enriched cell-process subnetworks in aorta samples from atherosclerotic NZW rabbits compared to control NZW rabbits. a: Analysis of cell processes related to overexpressed proteins in atherosclerotic plaque of hypercholesterolemic NZW rabbits. Major cell processes reported in Pathway Studio® as both functionally associated with atherosclerosis and regulated by multiple proteins with differential abundance in aorta protein extracts of atherosclerotic and healthy NZW rabbits are reported. Cell processes are grouped in 4 categories: immune-, vascular-, stress- and lipid-related events. All links are supported by biological findings from scientific literature. b: Enriched subnetworks related to lipid cell processes. Protein entities colorized in red and blue are respectively up- and down-regulated in aorta protein extracts of atherosclerotic compared to healthy NZW rabbits. Proteins highlighted in green are co-regulating in 3 or more of the 6 cell processes. All links are supported by biological findings from scientific literature. Variation of entity name police size does not have any biological meaning.

Figure 4. Implementation of the flow cytometry strategy. a: Control antibodies were tested by flow cytometry to guarantee the quality of the coupling strategy. AP2: mouse anti- α IIb β 3 antibody, CAII: rabbit anti-carbonic anhydrase II (CAII) antibody (Abcam) and RAM11: murine RAM11 antibody (anti-CD68). Secondary antibodies are Alexa 488 F(ab')₂ anti-mouse and Alexa 488 F(ab')₂ anti-rabbit. b: Fluorescence intensity (geomean) of hyperphage particles from 10⁸ to 10¹⁰ cfu/mL in A and D conditions.

Figure 5. FACS analysis typical experiment of atherosclerotic protein-binding scFv-phages. The number of coated beads (counts:Y-axis) is given as a function of the fluorescence intensity of scFv-phage staining (Alexa Fluor 488:X-axis). Protein binding was detected by mouse anti-pVIII antibody and Alexa Fluor 488-conjugated anti mouse antibody. The geomean of eight clones is given as a measure of the fluorescent intensity. The positivity is defined by at least 2-fold over the averaged fluorescence intensity of hyperphage particles at different concentrations.

Figure 6. Immunodetection of scFv-phages on sections of vulnerable atherosclerotic lesions from atherosclerotic NZW rabbits by immunohistochemical assays. Selected scFv-phages were incubated with atherosclerotic rabbit tissue sections. (a) Secondary antibody alone, (b) A2PL20, (c) B6PL20, (d) B7PL20, (e) C3PL20, (f) C9PL20, (g) C10PL20, (h) D3PL20, (i) E3PL20. After addition of secondary anti-pVIII protein antibody, sections were revealed with DAB substrate kit reagent. The presence of the antigen recognized by scFv-phages was indicated by a yellow-brown staining. No staining was observed in the presence of secondary antibody only. Scale bars represent 50 μ m. A 10X

magnification is presented for each scFv-phage. The picture on the right down corner corresponds to a magnification of 40X. Nuclei were counterstained with hematoxylin.

Tables

Gene Symbol	Gene Name	Ratio Atherosclerotic vs. Control*
FABP7	fatty acid binding protein 7, brain	253.7
FABP5	fatty acid binding protein 5 (psoriasis-associated)	202.5
ENO2	enolase 2 (gamma, neuronal)	182.2
ITGB2	integrin, beta 2 (complement component 3 receptor 3 and 4 subunit)	147.2
PLIN2	perilipin 2	131.9
ALB	albumin	106.8
CLU	clusterin	105.5
BAX	BCL2-associated X protein	101.8
PON3	paraoxonase 3	45.7
CCDC80	coiled-coil domain containing 80	38.4
GC	group-specific component (vitamin D binding protein)	38.0
PLG	plasminogen	37.7
HRG	histidine-rich glycoprotein	26.7
APMAP	adipocyte plasma membrane associated protein	21.9
A2M	alpha-2-macroglobulin	20.8
C4A	complement component 4A (Rodgers blood group)	19.8
ACOX1	acyl-CoA oxidase 1, palmitoyl	19.4
LTF	lactotransferrin	19.4
TF	transferrin	18.1
LGALS3	lectin, galactoside-binding, soluble, 3	17.7
APOA1	apolipoprotein A-I	17.3
HEXA	hexosaminidase A (alpha polypeptide)	17.1
CTBP1	C-terminal binding protein 1	16.8
CORO1A	coronin, actin binding protein, 1A	16.8
SCARB2	scavenger receptor class B, member 2	14.8
APOB	apolipoprotein B	14.5
CTSD	cathepsin D	14.1
ITGB3	integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)	13.7
AOX1	aldehyde oxidase 1	11.4
MFGE8	milk fat globule-EGF factor 8 protein	10.9
APOE	apolipoprotein E	10.8
HP	haptoglobin	10.1
COL1A1	collagen, type I, alpha 1	9.0
CP	ceruloplasmin (ferroxidase)	9.0
XDH	xanthine dehydrogenase	8.6

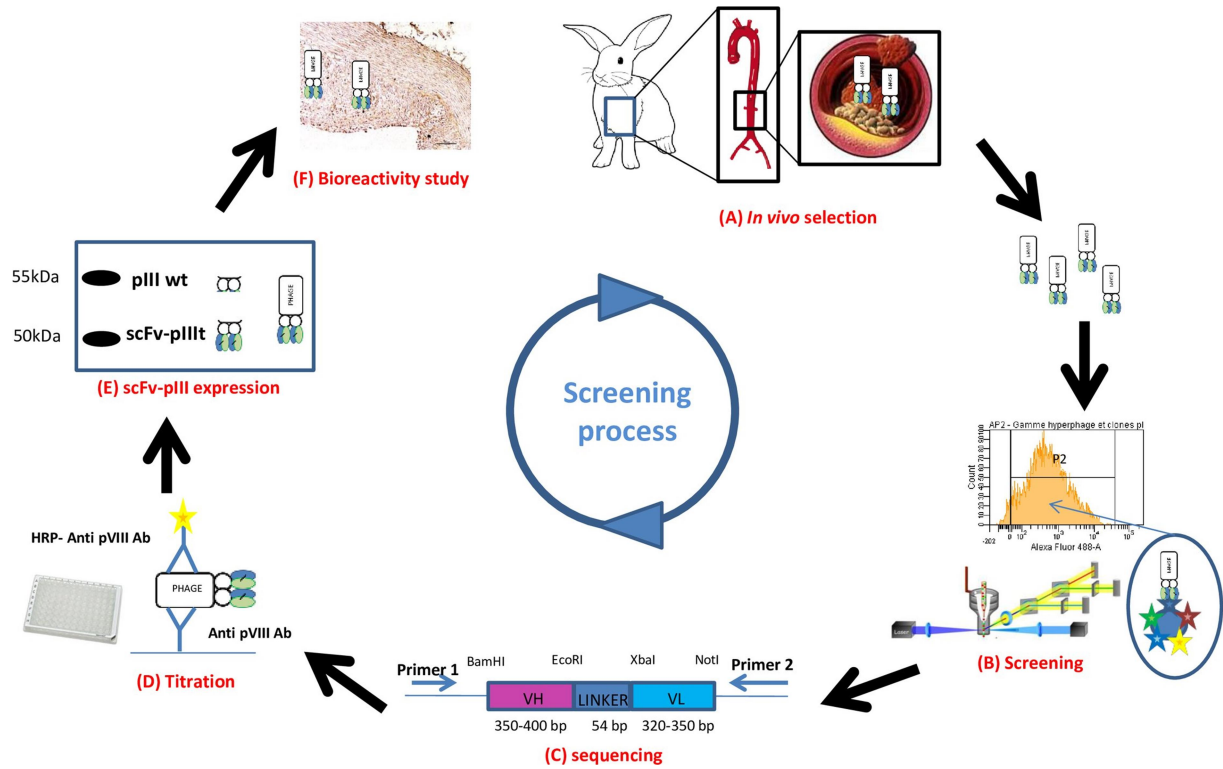
CA2	carbonic anhydrase II	8.4
CES2	carboxylesterase 2	8.0
CES1	carboxylesterase 1	6.5
UFM1	ubiquitin-fold modifier 1	6.3
CAT	catalase	5.8
	matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV	
MMP2	collagenase)	5.7
APP	amyloid beta (A4) precursor protein	4.7
AEBP1	AE binding protein 1	4.3
FN1	fibronectin 1	4.1
ACAT1	acetyl-CoA acetyltransferase 1	4.0
F2	coagulation factor II (thrombin)	3.7
FABP4	fatty acid binding protein 4, adipocyte	2.9
DPP4	dipeptidyl-peptidase 4	2.8
PRDX1	peroxiredoxin 1	2.8
DLD	dihydrolipoamide dehydrogenase	2.5
CALR	calreticulin	2.5
MAPK3	mitogen-activated protein kinase 3	2.2
GSR	glutathione reductase	2.2
CAV1	caveolin 1, caveolae protein, 22kDa	2.2
STAT1	signal transducer and activator of transcription 1, 91kDa	2.1
RHOA	ras homolog family member A	1.7
GRB2	growth factor receptor-bound protein 2	1.5
CD36	CD36 molecule (thrombospondin receptor)	1.5
LRP1	low density lipoprotein receptor-related protein 1	1.5
HSPA5	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	1.5
PRDX2	peroxiredoxin 2	1.4
PRKCD	protein kinase C, delta	1.4
	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP	
ATIC	cyclohydrolase	1.3
PRDX3	peroxiredoxin 3	1.2
CAPN1	calpain 1, (mu/I) large subunit	1.1
	adaptor protein, phosphotyrosine interaction, PH domain and leucine	
APPL1	zipper containing 1	1.1
SOD2	superoxide dismutase 2, mitochondrial	1.1
MAPK1	mitogen-activated protein kinase 1	1.1
PPIA	peptidylprolyl isomerase A (cyclophilin A)	1.1
ARF1	ADP-ribosylation factor 1	1.0
DNM2	dynamamin 2	0.9
APOA1		
BP	apolipoprotein A-I binding protein	0.9
RELA	v-rel avian reticuloendotheliosis viral oncogene homolog A	0.8
ROCK1	Rho-associated, coiled-coil containing protein kinase 1	0.8
STAT3	signal transducer and activator of transcription 3	0.8
VIM	vimentin	0.8
STK25	serine/threonine kinase 25	0.7

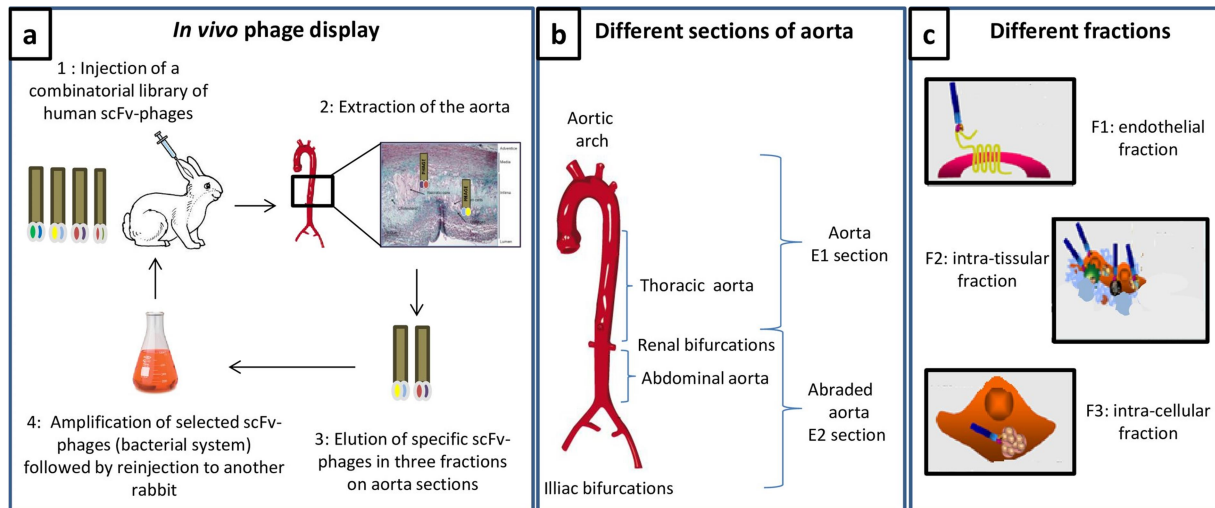
ROCK2	Rho-associated, coiled-coil containing protein kinase 2	0.7
GPD1	glycerol-3-phosphate dehydrogenase 1 (soluble)	0.7
GSK3B	glycogen synthase kinase 3 beta	0.7
GNAS	GNAS complex locus	0.6
HSPB1	heat shock 27kDa protein 1	0.5
GPX1	glutathione peroxidase 1	0.5
ACADM	acyl-CoA dehydrogenase, C-4 to C-12 straight chain	0.4
LSS	lanosterol synthase (2,3-oxidosqualene-lanosterol cyclase)	0.3
PLIN1	perilipin 1	0.2
PXDN	peroxidasin	0.0

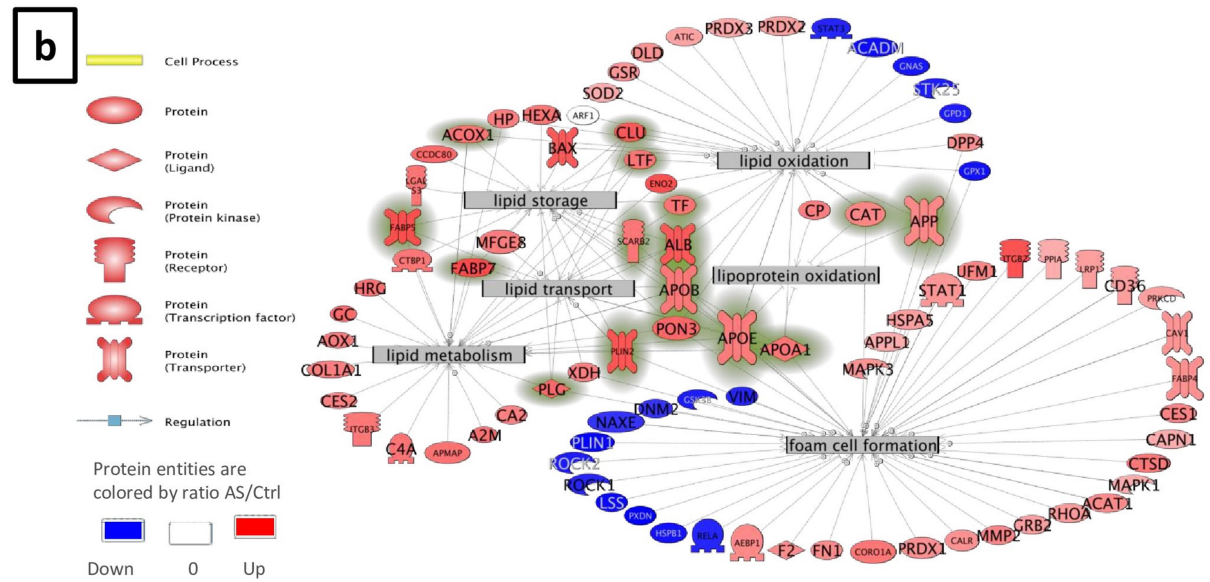
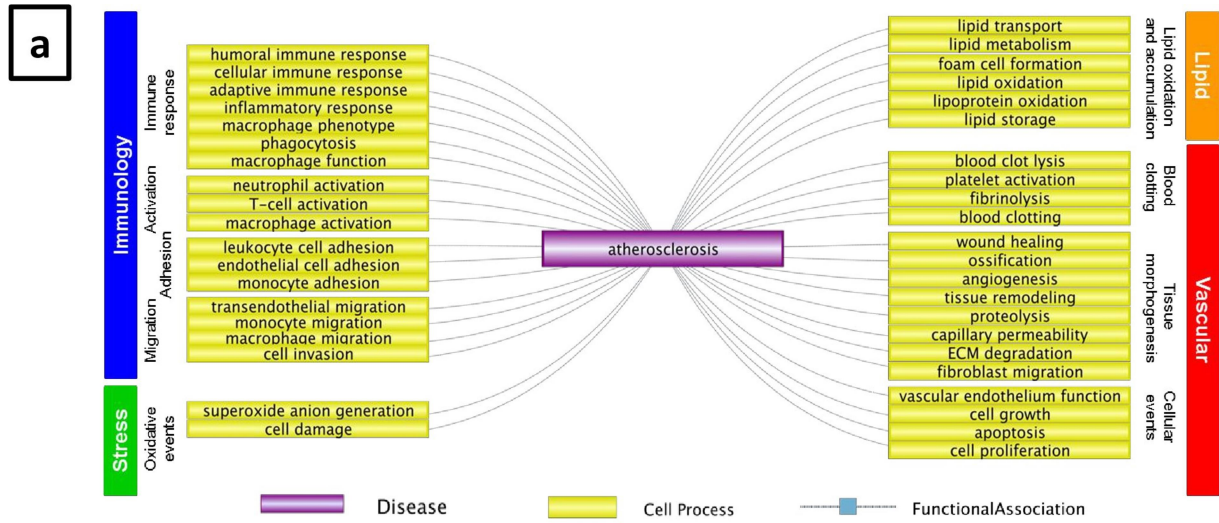
Table 1. Abundance of proteins belonging to lipid-related cell process subnetworks enriched in BIAHR rabbits compared to aorta of H rabbits. Protein entities colorized in red and blue are respectively up- and down-regulated in aorta samples of atherosclerotic compared to control NZW rabbits. Ratio values are reported.

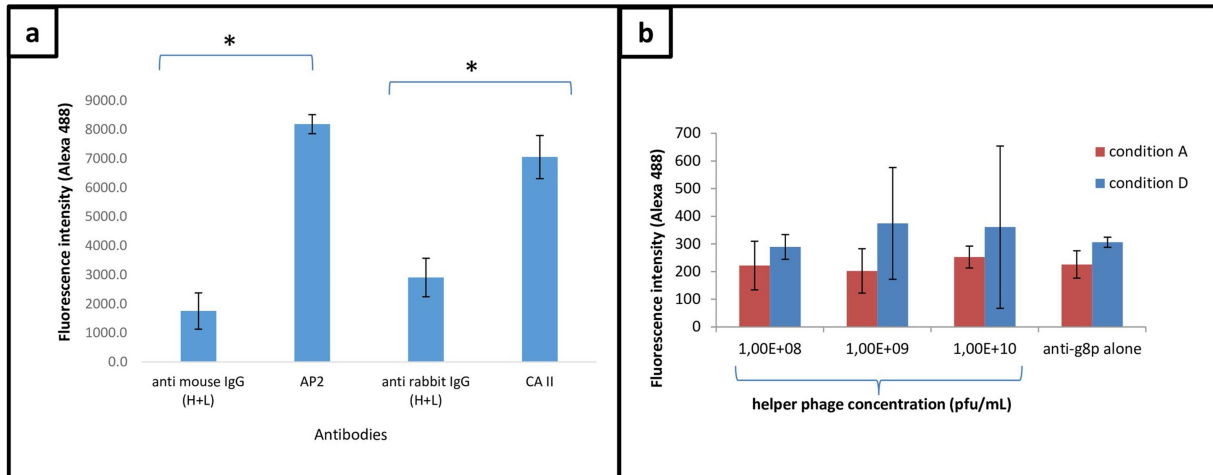
Conditions	Beads/phages incubation	Centrifugation	Washing	Anti pVIII incubation	Centrifugation	Washing	Secondary antibody incubation
Condition A	+	-	-	+	+	+	+
Condition B	+	-	-	+	+	-	+
Condition C	+	+	+	+	+	+	+
Condition D	+	+	-	+	-	-	+

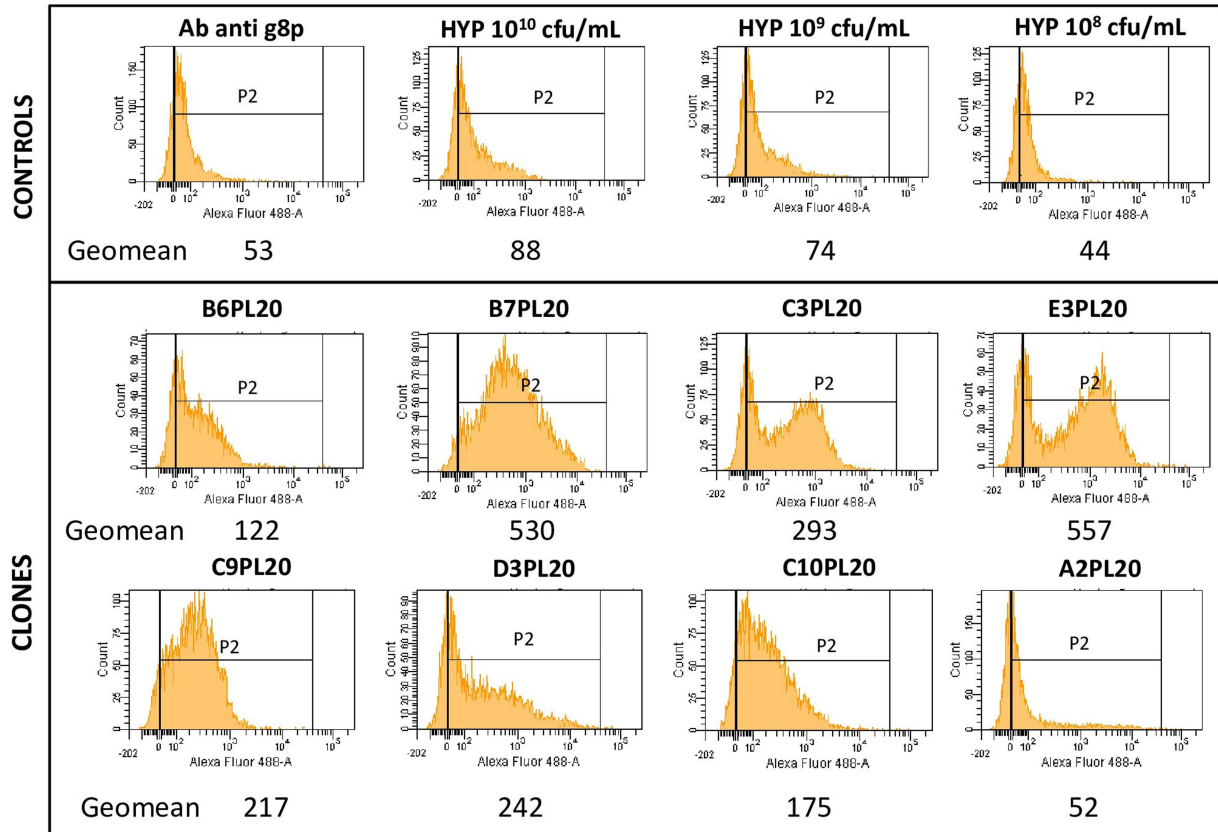
Table 2. Summary of the different conditions tested by FACS analysis to reduce the non-specific binding of hyperphage particles. The addition or not of the different steps is indicated by (+) or (-).

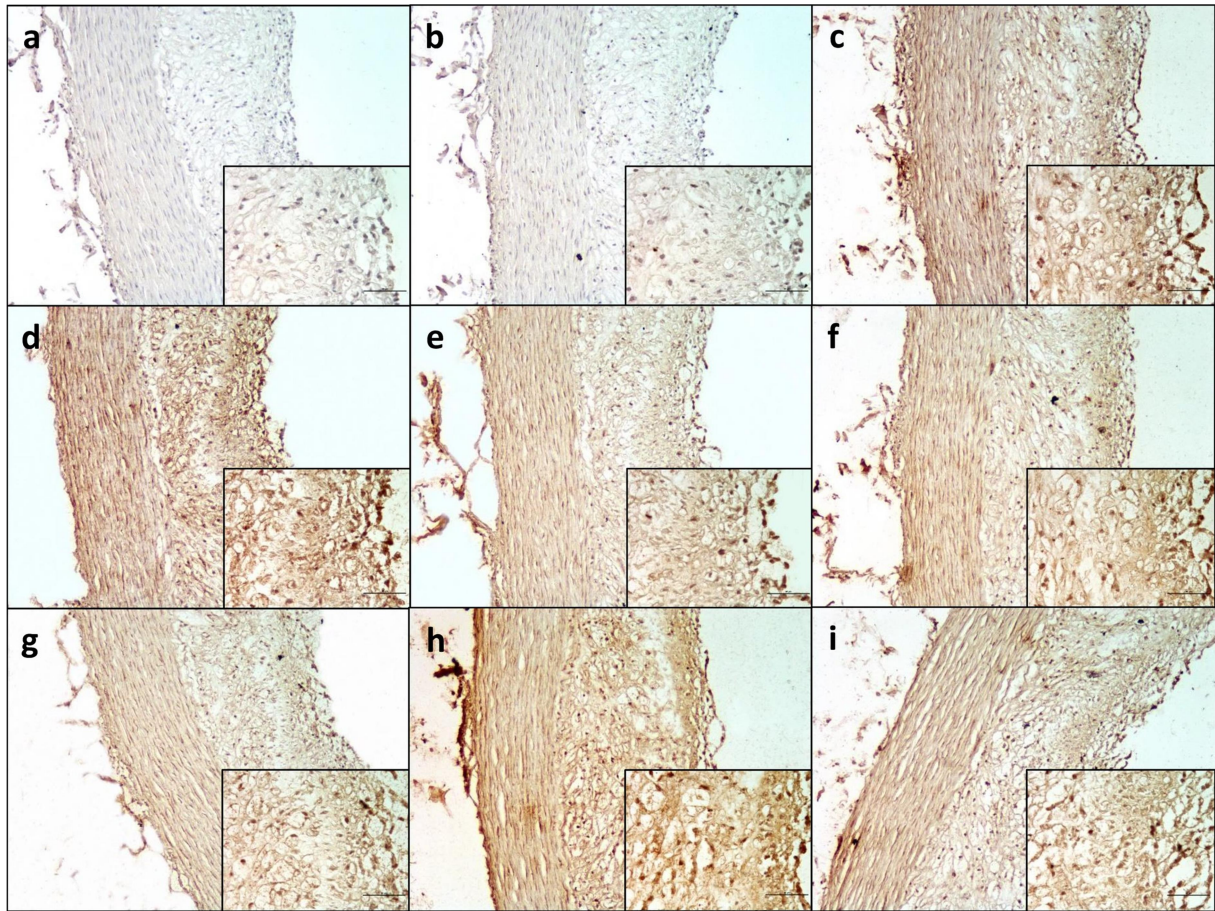












An innovative flow cytometry method for the screening of *in vivo* selected scFv-phages issuing from a fully human antibody library

Audrey HEMADOU¹, Jeanny LAROCHE-TRAINEAU¹, Ségolène ANTOINE¹, Philippe MONDON², Alexandre FONTAYNE², Yannick LE PRIOL³, Stéphane CLAVEROL⁴, Stéphane SANCHEZ¹, Martine CERUTTI⁵, Florence OTTONES¹, Gisèle CLOFENT-SANCHEZ¹ and Marie-Josée JACOBIN-VALAT^{†*1}

[†]Equivalent position of authors

¹CRMSB, UMR5536 CNRS, INSB, Bordeaux, 33076, France

²LFB Biotechnologies, department of biotherapeutic engineering Lille, 59000, France

³Elsevier Masson SAS, Elsevier RD solutions, Issy les Moulineaux, 92130, France

⁴CGFB, Proteome pole, Bordeaux, 33076, France

⁵UPS 3044, CNRS, Saint-Christol-Lès-Alès, France

1 Supplementary Methods

1.1 S1. ScFv sequencing

Sanger sequencing was performed using primers specific to flanking regions in the phagemide (5'-TGCAAATTCTATTTCAAGGAGAC-3' and 5'-AGAATCATCAGATAAAGTAATCC-3'). Antibody gene fragments were analyzed by IMGT/V-QUEST database (www.imgt.org/IMGT_vquest/vquest) for V germline determination and CDR3 analyses.

1.2 S2. Western blot analyses of scFv displayed on phages

After migration of the scFv-phages on 4-15% SDS-PAGE followed by western blotting, the scFv expression was detected using as primary mouse anti-pIII protein antibody (Mobitec, Ozyme, France) at 1 µg/mL in MPBS 0.5% as described earlier¹². Secondary anti mouse HRP (1:1000) antibody was added and incubated at RT for 2 h. Each step was followed by extensive washing and chemoluminescent signal was developed with ECL reagent (ThermoFisher, France).

1.3 S3. Immunohistochemical analysis on NZW rabbit sections

Clones screened positively by flow cytometry were evaluated for binding to paraffin-embedded atheromatous tissue samples prepared from the same atheromatous specimen. Briefly, after a retrieval step in retrieval buffer (TRIS 10mM, EDTA 1mM, 0.05% Tween 20, pH 9), deparaffinized tissue sections were incubated with scFv-phages overnight at 4°C, washed with PBS BSA 2%. Sections were then incubated with mouse anti-pVIII protein antibody diluted at 1:200 (Abcam, France) for 1h and then with envision HRP system anti mouse. Binding was detected using DAB system (Dako, France).

1.4 S4. In vivo phage-display

The elution protocols of endothelial, intra-tissular and intra-cellular scFv fractions were performed according to Deramchia et al⁷, with minor modifications. The endothelium cell surface-bound scFv-phage fraction (F1 fraction) was eluted with 500 µL of 0.1 M glycine-HCl, pH 2.2 and immediately neutralized with 15 µL of 2.5 M Tris-HCl, pH 8. The elution procedure was repeated and the different samples were pooled, a protease-inhibitor cocktail was added (ThermoScientific) and the fraction was stored at 4 °C prior bacterial infection.

In order to elute intra-tissular scFv-phage fraction (F2 fraction), the aortic tissue was incubated with 900 μL of PBS (Ca^{2+} , Mg^{2+} free) containing 2,000 U/mL of collagenase type II (Gibco) adjusted to 1 mL with 2.5% Trypsin-EDTA (Eurobio) and scratched using a glass slide to facilitate the tissue dissociation. Cycles of 20 s Polytron homogenizer (Ultraturax TP-20, Kinematica) were then performed at 4 °C to obtain a homogeneous solution. The homogenate was then incubated at 37 °C for 30 min with punctual vortexing and finally centrifuged for 10 minutes at 1,000 g to remove insoluble material. This preparation was homogenized 2 times more following the same procedure. After each centrifugation, supernatants containing eluted scFv-phages were collected and pooled in a clean tube in presence of a protease-inhibitor cocktail.

To access the internalized scFv-phage fraction (F3 fraction), the insoluble material was incubated with 500 μL of 0.1 M TEA (Tetraethylammonium chloride) (Sigma-Aldrich), pH 11.5 for 5 min at RT and then vigorously vortexed for another 5 min. Samples were neutralized by addition of 150 μL of 1M Tris-HCl, pH 7.4. After centrifugation at 1,000 g for 10 min, the supernatant was collected.

2 Supplementary Figure legends

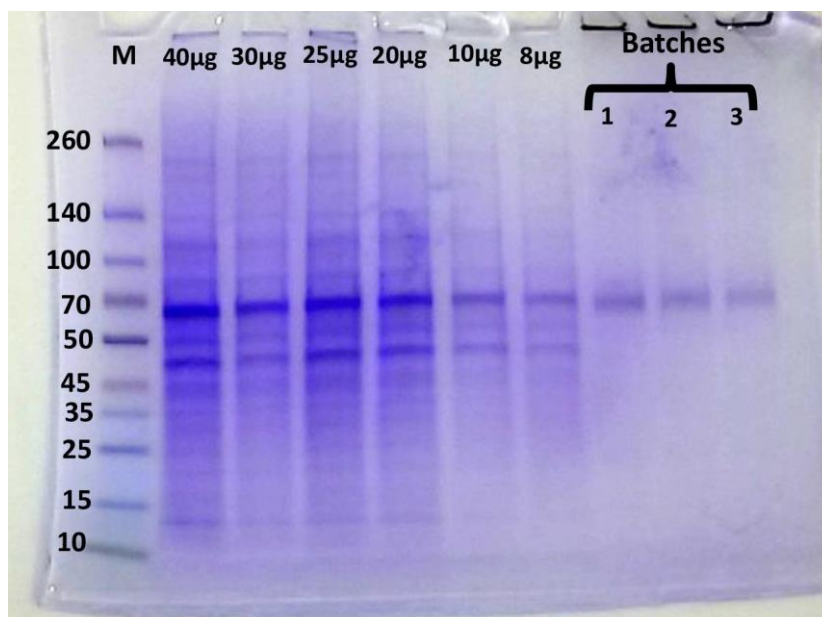
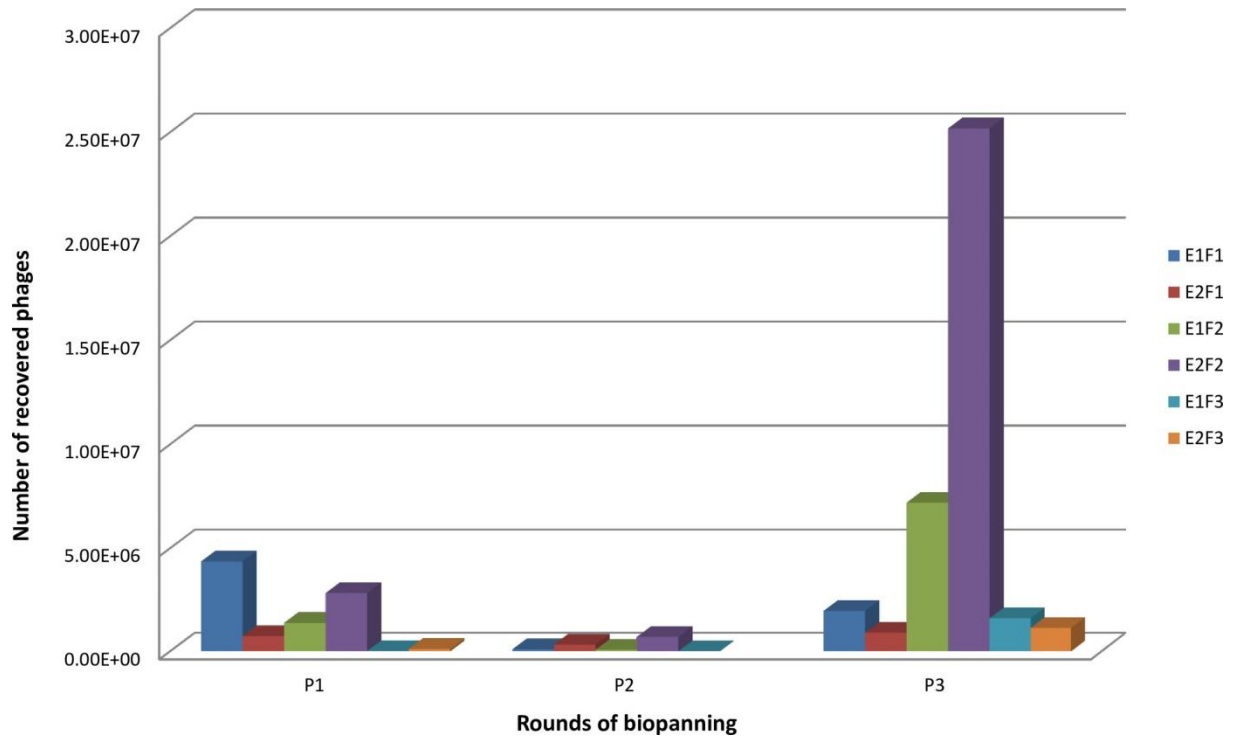
Figure S1. Phage enrichment after three rounds of *in vivo* phage display biopanning. P1: first round of biopanning, P2: second round and P3: third round. E1: aortic arch and thoracic area, E2: abraded renal and abdominal areas. F1: endothelial fraction, F2: intra-tissular fraction, F3: intra-cellular fraction.

Figure S2. Evaluation of the coupling rate of BIAHR proteins on beads. A range of BIAHR proteins (40 to 8 μg) was used to evaluate the coupling rate of proteins on beads by SDS-PAGE. Image J analysis allowed to estimate the coupling rate at 80%.

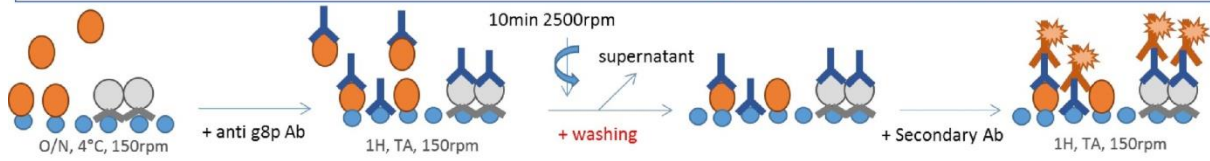
Figure S3. Schematic representation of A and D conditions tested for the implementation of the flow cytometry strategy.

Figure S4. ScFv-phage production analysed by SDS-PAGE. Example of eight scFv-phage clones produced after infection of recombinant bacteria with hyperphage. Proteins were separated on 4-15% Bis acrylamide gel. The pIII protein fragment was detected by mouse anti-pIII protein antibody and HRP anti mouse antibody. SuperSignal west pico chemiluminescent substrate was used for staining the membrane. Images were analysed by Snapgene software. The white lineage corresponds to a crop of the western blot membrane.

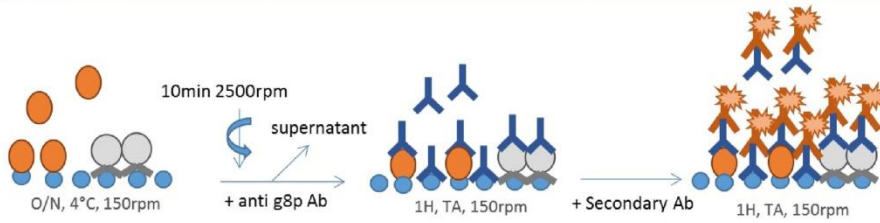
3 Supplementary figures



Condition A: beads + scFv-phages + anti g8p Ab → centrifugation + **washing** → Labeled secondary Ab (anti mouse)

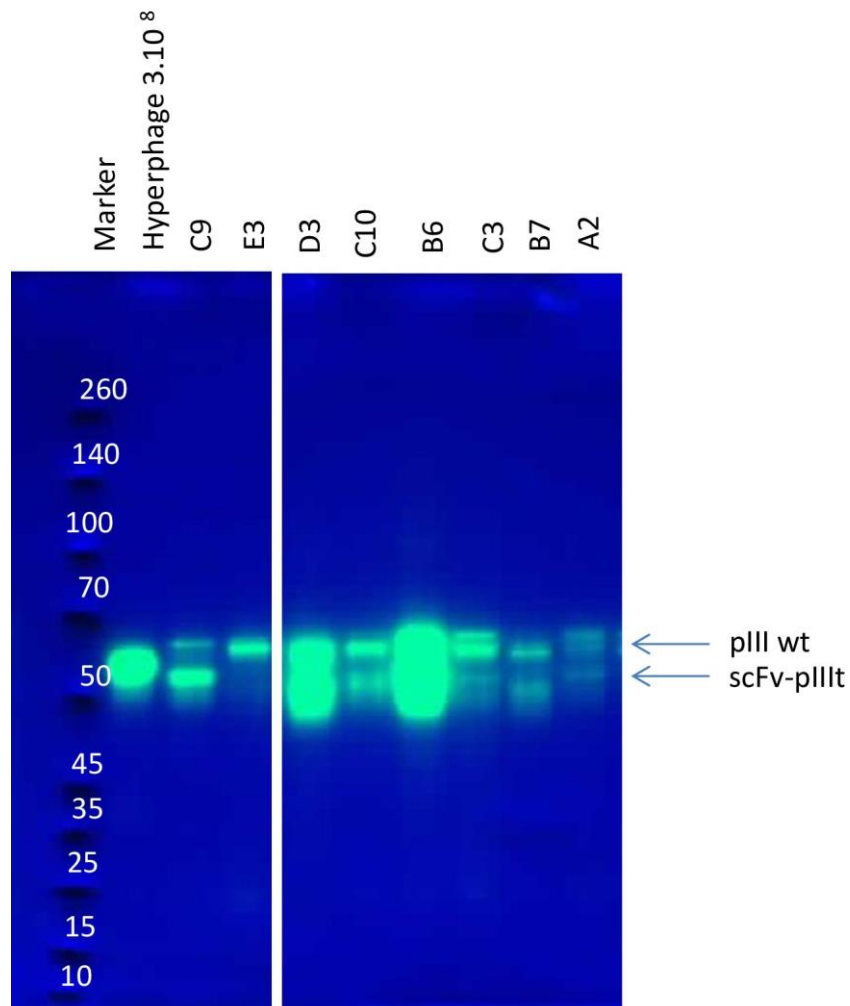


Condition D: beads + scFv-phages → centrifugation → anti g8p Ab → Labeled secondary Ab (anti mouse)



Legend

- Helper phage
- Y Ab anti g8p
- Beads/atheromatous proteins
- scFv-phages
- ★ Alexa 488 Secondary Ab



4 Supplementary Tables

ratio proteins	BIAHR/H	Description
7898.19		Fibulin-7 OS=Homo sapiens GN=FBLN7 PE=2 SV=1 - [FBLN7_HUMAN]
3478.34		Lysophosphatidylcholine acyltransferase 2 OS=Homo sapiens GN=LPCAT2 PE=1 SV=1 - [PCAT2_HUMAN]
253.68		Fatty acid-binding protein, brain OS=Homo sapiens GN=FABP7 PE=1 SV=3 - [FABP7_HUMAN]
234.64		Beta-galactosidase-1-like protein OS=Homo sapiens GN=GLB1L PE=2 SV=1 - [GLB1L_HUMAN]
215.31		Receptor-type tyrosine-protein phosphatase C OS=Homo sapiens GN=PTPRC PE=1 SV=2 - [PTPRC_HUMAN]
202.46		Fatty acid-binding protein, epidermal OS=Homo sapiens GN=FABP5 PE=1 SV=3 - [FABP5_HUMAN]
182.25		Gamma-enolase OS=Homo sapiens GN=ENO2 PE=1 SV=3 - [ENOG_HUMAN]
147.57		Polyadenylate-binding protein 4 OS=Homo sapiens GN=PABPC4 PE=1 SV=1 - [PABP4_HUMAN]
147.16		Integrin beta-2 OS=Homo sapiens GN=ITGB2 PE=1 SV=2 - [ITB2_HUMAN]
131.89		Perilipin-2 OS=Homo sapiens GN=PLIN2 PE=1 SV=2 - [PLIN2_HUMAN]
131.69		Hypothetical drug-resistance-associated protein OS=Homo sapiens PE=4 SV=1 - [Q86YX8_HUMAN]
122.44		Carabin OS=Homo sapiens GN=TBC1D10C PE=1 SV=1 - [TB10C_HUMAN]
109.92		Alpha-actinin-1 OS=Homo sapiens GN=ACTN1 PE=1 SV=2 - [ACTN1_HUMAN]
106.83		Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [ALBU_HUMAN]
105.52		Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1 - [CLUS_HUMAN]
101.81		Apoptosis regulator BAX OS=Homo sapiens GN=BAX PE=1 SV=1 - [BAX_HUMAN]
79.10		Tetranectin OS=Homo sapiens GN=CLEC3B PE=1 SV=3 - [TETN_HUMAN]
76.12		Transmembrane channel-like protein 2 OS=Homo sapiens GN=TMC2 PE=2 SV=3 - [TMC2_HUMAN]
71.83		Cartilage oligomeric matrix protein OS=Homo sapiens GN=COMP PE=1 SV=2 - [COMP_HUMAN]
64.54		Ras-related C3 botulinum toxin substrate 2 OS=Homo sapiens GN=RAC2 PE=1 SV=1 - [RAC2_HUMAN]
62.99		Probable ATP-dependent RNA helicase DDX6 OS=Homo sapiens GN=DDX6 PE=1 SV=2 - [DDX6_HUMAN]
62.63		Syntenin-1 OS=Homo sapiens GN=SDCBP PE=1 SV=1 - [SDCB1_HUMAN]
57.28		Beta-glucuronidase OS=Homo sapiens GN=GUSB PE=1 SV=2 - [BGLR_HUMAN]
54.16		Serine protease HTRA1 OS=Homo sapiens GN=HTRA1 PE=1 SV=1 - [HTRA1_HUMAN]
50.26		Macrophage-capping protein OS=Homo sapiens GN=CAPG PE=1 SV=2 - [CAPG_HUMAN]
48.10		Histamine N-methyltransferase OS=Homo sapiens GN=HNMT PE=1 SV=1 - [HNMT_HUMAN]
45.74		Serum paraoxonase/lactonase 3 OS=Homo sapiens GN=PON3 PE=1 SV=3 - [PON3_HUMAN]
45.43		Purine nucleoside phosphorylase OS=Homo sapiens GN=PNP PE=1 SV=2 -

	[PNPH_HUMAN]
42.24	Inter-alpha-trypsin inhibitor heavy chain H1 OS=Homo sapiens GN=ITIH1 PE=1 SV=3 - [ITIH1_HUMAN]
41.21	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1 - [IGHG1_HUMAN]
38.53	Alpha-N-acetylgalactosaminidase OS=Homo sapiens GN=NAGA PE=1 SV=2 - [NAGAB_HUMAN]
38.38	Coiled-coil domain-containing protein 80 OS=Homo sapiens GN=CCDC80 PE=1 SV=1 - [CCD80_HUMAN]
38.05	Vitamin D-binding protein OS=Homo sapiens GN=GC PE=1 SV=1 - [VTDB_HUMAN]
37.69	Plasminogen OS=Homo sapiens GN=PLG PE=1 SV=2 - [PLMN_HUMAN]
35.82	Protein TANC1 OS=Homo sapiens GN=TANC1 PE=1 SV=3 - [TANC1_HUMAN]
31.59	Complement component C7 OS=Homo sapiens GN=C7 PE=1 SV=2 - [CO7_HUMAN]
31.15	UPF0554 protein C2orf43 OS=Homo sapiens GN=C2orf43 PE=1 SV=1 - [CB043_HUMAN]
26.66	Histidine-rich glycoprotein OS=Homo sapiens GN=HRG PE=1 SV=1 - [HRG_HUMAN]
25.58	60S ribosomal protein L35a OS=Homo sapiens GN=RPL35A PE=1 SV=2 - [RL35A_HUMAN]
22.74	Plastin-2 OS=Homo sapiens GN=LCP1 PE=1 SV=6 - [PLSL_HUMAN]
21.94	Adipocyte plasma membrane-associated protein OS=Homo sapiens GN=APMAP PE=1 SV=2 - [APMAP_HUMAN]
20.85	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=3 - [A2MG_HUMAN]
20.41	Retinal dehydrogenase 2 OS=Homo sapiens GN=ALDH1A2 PE=1 SV=3 - [AL1A2_HUMAN]
19.84	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=2 - [CO4A_HUMAN]
19.81	Insulin-like growth factor-binding protein 7 OS=Homo sapiens GN=IGFBP7 PE=1 SV=1 - [IBP7_HUMAN]
19.43	Peroxisomal acyl-coenzyme A oxidase 1 OS=Homo sapiens GN=ACOX1 PE=1 SV=3 - [ACOX1_HUMAN]
19.42	Ferritin heavy chain OS=Homo sapiens GN=FTH1 PE=1 SV=2 - [FRIH_HUMAN]
19.39	Lactotransferrin OS=Homo sapiens GN=LTF PE=1 SV=6 - [TRFL_HUMAN]
19.36	Glypican-6 OS=Homo sapiens GN=GPC6 PE=1 SV=1 - [GPC6_HUMAN]
18.69	Pro-cathepsin H OS=Homo sapiens GN=CTSH PE=1 SV=4 - [CATH_HUMAN]
18.53	Complement C5 OS=Homo sapiens GN=C5 PE=1 SV=4 - [CO5_HUMAN]
18.06	Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=3 - [TRFE_HUMAN]
17.91	60S ribosomal protein L36a OS=Homo sapiens GN=RPL36A PE=1 SV=2 - [RL36A_HUMAN]
17.69	Galectin-3 OS=Homo sapiens GN=LGALS3 PE=1 SV=5 - [LEG3_HUMAN]
17.68	Protein flightless-1 homolog OS=Homo sapiens GN=FLII PE=1 SV=2 - [FLII_HUMAN]
17.54	2,4-dienoyl-CoA reductase, mitochondrial OS=Homo sapiens GN=DEC1 PE=1 SV=1 - [DEC1_HUMAN]
17.33	Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1 -

	[APOA1_HUMAN]
17,11	Beta-hexosaminidase subunit alpha OS=Homo sapiens GN=HEXA PE=1 SV=2 - [HEXA_HUMAN]
16.78	C-terminal-binding protein 1 OS=Homo sapiens GN=CTBP1 PE=1 SV=2 - [CTBP1_HUMAN]
16.76	Coronin-1A OS=Homo sapiens GN=CORO1A PE=1 SV=4 - [COR1A_HUMAN]
16.49	Complement component C6 OS=Homo sapiens GN=C6 PE=1 SV=3 - [CO6_HUMAN]
15.75	3-hydroxybutyrate dehydrogenase type 2 OS=Homo sapiens GN=BDH2 PE=1 SV=2 - [BDH2_HUMAN]
15.64	Twinfilin-1 OS=Homo sapiens GN=TWF1 PE=1 SV=3 - [TWF1_HUMAN]
15.59	Glycylpeptide N-tetradecanoyltransferase 1 OS=Homo sapiens GN=NMT1 PE=1 SV=2 - [NMT1_HUMAN]
14.75	Lysosome membrane protein 2 OS=Homo sapiens GN=SCARB2 PE=1 SV=2 - [SCRB2_HUMAN]
14.71	Galectin-3-binding protein OS=Homo sapiens GN=LGALS3BP PE=1 SV=1 - [LG3BP_HUMAN]
14.53	Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=2 - [APOB_HUMAN]
14.48	Ig kappa chain V-I region Roy OS=Homo sapiens PE=1 SV=1 - [KV116_HUMAN]
14.42	Collagen triple helix repeat-containing protein 1 OS=Homo sapiens GN=CTHRC1 PE=1 SV=1 - [CTHR1_HUMAN]
14.08	Cathepsin D OS=Homo sapiens GN=CTSD PE=1 SV=1 - [CATD_HUMAN]
14.06	Fibrinogen gamma chain OS=Homo sapiens GN=FGG PE=1 SV=3 - [FIBG_HUMAN]
13.69	Integrin beta-3 OS=Homo sapiens GN=ITGB3 PE=1 SV=2 - [ITB3_HUMAN]
13.44	Hemopexin OS=Homo sapiens GN=HPX PE=1 SV=2 - [HEMO_HUMAN]
13.34	Glia-derived nexin OS=Homo sapiens GN=SERPINE2 PE=1 SV=1 - [GDN_HUMAN]
13.02	Flotillin-2 OS=Homo sapiens GN=FLOT2 PE=1 SV=2 - [FLOT2_HUMAN]
12.96	Eukaryotic translation initiation factor 4 gamma 1 OS=Homo sapiens GN=EIF4G1 PE=2 SV=1 - [E7EX73_HUMAN]
12.75	Protein FAM49B OS=Homo sapiens GN=FAM49B PE=1 SV=1 - [FA49B_HUMAN]
12.56	EH domain-containing protein 4 OS=Homo sapiens GN=EHD4 PE=1 SV=1 - [EHD4_HUMAN]
12.42	Beta-galactosidase OS=Homo sapiens GN=GLB1 PE=1 SV=2 - [BGAL_HUMAN]
11.96	Protein S100-A4 OS=Homo sapiens GN=S100A4 PE=1 SV=1 - [S10A4_HUMAN]
11.78	Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2 PE=1 SV=3 - [IGHA2_HUMAN]
11.40	Aldehyde oxidase OS=Homo sapiens GN=AOX1 PE=1 SV=2 - [AOXA_HUMAN]
11.37	Alanine--tRNA ligase, cytoplasmic OS=Homo sapiens GN=AARS PE=1 SV=2 - [SYAC_HUMAN]
11.26	Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2 - [FIBB_HUMAN]
11.18	Soluble scavenger receptor cysteine-rich domain-containing protein SSC5D OS=Homo sapiens GN=SSC5D PE=2 SV=3 - [SRCRL_HUMAN]
11.06	Twinfilin-2 OS=Homo sapiens GN=TWF2 PE=1 SV=2 - [TWF2_HUMAN]
10.91	Lactadherin OS=Homo sapiens GN=MFG8 PE=1 SV=2 - [MFGM_HUMAN]

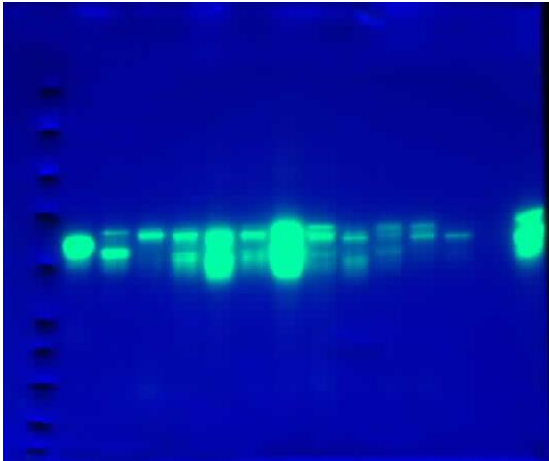
10.83	Transmembrane protein 245 OS=Homo sapiens GN=TMEM245 PE=1 SV=2 - [TM245_HUMAN]
10.82	Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1 - [APOE_HUMAN]
10.63	b(0,+)-type amino acid transporter 1 OS=Homo sapiens GN=SLC7A9 PE=4 SV=1 - [K7EKDO_HUMAN]
10.55	Glia maturation factor gamma OS=Homo sapiens GN=GMFG PE=1 SV=1 - [GMFG_HUMAN]
10.13	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1 - [HPT_HUMAN]
9.94	Fibulin-1 OS=Homo sapiens GN=FBLN1 PE=1 SV=4 - [FBLN1_HUMAN]
9.85	Dehydrogenase/reductase SDR family member 4 OS=Homo sapiens GN=DHRS4 PE=1 SV=3 - [DHRS4_HUMAN]
9.29	2-oxoisovalerate dehydrogenase subunit alpha, mitochondrial OS=Homo sapiens GN=BCKDHA PE=1 SV=2 - [ODBA_HUMAN]
9.01	Collagen alpha-1(I) chain OS=Homo sapiens GN=COL1A1 PE=1 SV=5 - [CO1A1_HUMAN]
8.97	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1 - [CERU_HUMAN]
8.95	Disabled homolog 2 OS=Homo sapiens GN=DAB2 PE=1 SV=3 - [DAB2_HUMAN]
8.64	Xanthine dehydrogenase/oxidase OS=Homo sapiens GN=XDH PE=1 SV=4 - [XDH_HUMAN]
8.40	V-type proton ATPase subunit E 1 OS=Homo sapiens GN=ATP6V1E1 PE=1 SV=1 - [VATE1_HUMAN]
8.37	Carbonic anhydrase 2 OS=Homo sapiens GN=CA2 PE=1 SV=2 - [CAH2_HUMAN]

Table S2. Ratio of proteins in BIAHR extract.

Masterblock	Fractions	Selected clones	Clones with an in frame scFv sequence	Nb of VH clonotypes	Nb of VL clonotypes
PL19	E1F1	12	8	7	8
PL20	E2F1	27	17	15	17
PL21	E1F2	15	11	10	11
PL22	E1F2	36	19	19	18
PL23	E2F2	61	45	40	41
PL24	E2F2	10	9	9	9
PL25	E1F3	13	11	11	10
PL26	E1F3	18	10	10	10
PL27	E2F3	17	12	10	10

Table S3. Sanger Sequencing and IMGT results on 209 clones selected by flow cytometry. Among 209 clones selected by flow cytometry, 68% of them have a complete sequence with 42% without stop codons. Analyses of the germlines and CDR3 junctions showed a high diversity among the clones.

Additional Information – supplementary data



An innovative flow cytometry method for the screening of *in vivo* selected scFv-phages issuing from a fully human antibody library

Audrey HEMADOU¹, Jeanny LAROCHE-TRAINEAU^{†1}, Ségolène ANTOINE^{†1}, Philippe MONDON², Alexandre FONTAYNE², Yannick LE PRIOL³, Stéphane CLAVEROL⁴, Stéphane SANCHEZ¹, Martine CERUTTI⁵, Florence OTTONES¹, Gisèle CLOFENT-SANCHEZ^{†1} and Marie-Josée JACOBIN-VALAT^{†*1}

[†]Equivalent position of authors

¹ CRMSB, UMR5536 CNRS, INSB, Bordeaux, 33076, France

² LFB Biotechnologies, department of biotherapeutic engineering Lille, 59000, France

³ Elsevier Masson SAS, Elsevier RD solutions, Issy les Moulineaux, 92130, France

⁴ CGFB, Proteome pole, Bordeaux, 33076, France

⁵ UPS 3044, CNRS, Saint-Christol-Lès-Alès, France

Gene Set Seed	Category	TOP100 SNEA	SNEA p-value	Measured Entities	TOP100 FSNE	FSNE p-value	Overlapping Entities	Cell Process Linked to atherosclerosis in Pathway Studio	Display Figure 1
transendothelial migration	Immunology	N	-	-	Y	1.5E-10	PTPRC;LGALS3;LGALS3BP;HTRA1;ITGB2;APOB;RAC2;APOA1;PLG;HP;ALB;XDH;A2M;S100A4;ITGB3;FGF	Y	Y
T-cell activation	Immunology	N	-	-	Y	8.6E-11	PTPRC;LGALS3;LGALS3BP;ITGB2;BAX;APOB;LTF;LCP1;FLOT2;HRG;APOA1;MFGE8;RAC2;HP;ALB;ACTN1;TF;HPRT1;S100A4;APOE;FTH1;CORO1A	Y	Y
phagocytosis	Immunology	N	-	-	Y	2.5E-15	PABPC4;GMFG;C5;LTF;HRG;RAC2;MFGE8;PLG;HP;ALB;C4A;COL1A1;CP;TF;APOE;FABP7;CLU;CAPG;CORO1A;PTPRC;LGALS3;ITGB2;LCP1;APOA1;GUSB;SCARB2;A2M;HPX;ITGB3;FTH1	Y	Y
neutrophil activation	Immunology	Y	0.0103	1;TLN1;RAP1A;MARKS;CRKL	Y	1.0E-13	PTPRC;LGALS3;LGALS3BP;ITGB2;BAX;APOB;C5;LTF;LCP1;APOA1;RAC2;HP;ALB;ACTN1;C4A;XDH;A2M;APOE	Y	Y

monocyte migration	Y	0.0041	Y	N	-	-	ENO2;ITGB2;ACTN1;ALB;CLU;RAC2;PLG;LTF;LGA LS3;GMFG;HP;CAT;HSP2;APP;TXN;FN1;F 2;GLG1;SOD3;ENO1;GPI;DPP4;PRDX1;ANXA2;SE RPINF1;CALR;MAPK3;CAV1;STAT1;TGFB1;RAC1; ANXA1;RHOA;CD9;CDC42;LRP1;CD36;YWHAZ;IT GA5;RPS19;MAPK1;PIIA;HMGB1;LAMC1;ITGAL;R OCK1;AOC3;VIM;GSK3B;ITGB1;GNAQ;HSPB1;TE S;TGM2;MYLK;PTK2;DYSF;ELN	Y	Y	Y
monocyte adhesion	N	-	Y	Y	1.7E-08	-	PTPRC;LGALS3;ITGB2;LGALS3BP;GMFG; APOB;LTF;APOA1;PLG;SERPINE2;ALB;XD H;APOE;ITGB3	Y	Y	Y
macrophage phenotype	Y	0.0089	Y	Y	4.9E-10	-	LGALS3;COMP;HPX;APOE;A2M;FABP5;M FGE8;APOA1;RAC2;HP;ACTN1	Y	Y	Y
macrophage migration	Y	0.0166	Y	Y	1.3E-09	-	PTPRC;DAB2;LGALS3;ITGB2;LTF;RAC2;A POA1;PLG;A2M;CLU;ITGB3;ACOX1;FTH1	Y	Y	Y
macrophage function	Y	0.0008	Y	Y	3.2E-13	-	PTPRC;LGALS3;ITGB2;LGALS3BP;C5;LTF; RAC2;APOA1;PLG;HP;ACTN1;ALB;TF;A2M ;APOE;FABP5;CLU;CAPG;CORO1A	Y	Y	Y
macrophage activation leukocyte cell adhesion	Y	0.0201	Y	Y	6.0E-11	-	PTPRC;LGALS3;ITGB2;LTF;HRG;APOA1;P LG;HP;ALB;C4A;FLI1;TF;A2M;APOE;FABP5 ;CORO1A;GC	Y	Y	Y
	N	-	Y	Y	7.1E-09	-	PTPRC;LGALS3;ITGB2;COMP;LTF;LCP1;A POA1;PLG;HP;ALB;XDH;CP;TF;A2M;ITGB3	Y	Y	Y

inflammatory response	Immunology	N	-	0.0068	Y	1.3E-16	GC	C6:APOB;LTF;C5;HRG;MFGE8;RAC2;PLG;HP;ALB;C4A;CTSD;FLII;XDH;CP;CSFB;FG;IGFBP7;TF;APOE;FABP5;CLU;CAPG;CORO1A;PTPRC;DAB2;LPCAT2;LGALS3;ITGB2;BAX;LCP1;APOA1;SERPINE2;GUSB;C7;CTSH;SCARB2;A2M;S100A4;HPX;ITGB3;
humoral immune response	Immunology	Y	0.0068	0.0102	Y	-	N	PTPRC;ALB;LCP1;A2M;ALDH1A2;C4A;LTF;TF;LGALS3;APOA1;LGALS3BP;APOE;FBLN1;VTN;TNFRSF6B;CAT;PAD2;APP;SPAG9;TXN;MPZ;FNI;F2;SO;D3;GPI;FABP4;NQO1;DPP4;PRDX1;NPM1;TPI1;FO;LH1;CALR;RPL5;IDH1;SET;ANXA1;RPL23;STIP1;RPLP0;EIF4A3;CDC42;HSPA5;YWHAQ;EIF4A1;HNRNP2B1;HSPD1;EEF2;HSP90B1;CYBB;GLS;MAPK1;HMGB1;HNRNP1;EEF1A1;RAP1B;C3;STAT3;IDH2;TGM2;RPL12
cellular immune response	Immunology	Y	0.0102	0.0007	Y	-	N	PTPRC;ITGB2;ALB;CLU;BAX;RAC2;PNP;LCP1;A2M;C4A;LTF;LGALS3;LGALS3BP;APOE;TNFRSF6B;CAT;HSPE1;MVP;APP;TXN;HSP90AA1;G6PD;HLA-DQB1;HLA-B;F2;FKBP1A;ENO1;GPI;PSME2;NQO1;DPP4;GSN;CALR;MAPK3;GSR;STAT1;RAC1;RPL23;RHOA;RPL7;P4HB;HSPA8;CDC42;LRP1;CD36;GAPDH;EIF4A1;CTSC;CRYAB;HSPD1;PARK7;HSP90B1;SOD2;MAPK1;PIPA;HSPB6;HMGB1;NDUFA13;EEF1A1;RPN1;HSPA2;RELA;STAT3;RPSA;PTPN11;ACTB;CSPG4;JUP;GPX1;DCN;HSPB1;TES;TGM2;RPL12;ATP2A3;DMD;NRP1;ELN
adaptive immune response	Immunology	Y	0.0185	0.0007	Y	3.1E-13	Y	PTPRC;ITGB2;ALB;CLU;RAC2;LCP1;A2M;ALDH1A2;C4A;LTF;C5;LGALS3;APOA1;CORO1A;C6;LGALS3BP;APOB;MFGE8;APOE;HP;XDH;TNFRSF6B;CAT;HSPE1;APP;MPZ;FNI;HLA-B;F2;SOD3;ENO1;DPP4;CALR;HBB;MAPK3;STAT1;ANXA1;HSPA8;LRP1;CD36;UBE2L3;CTSC;PHB2;BGN;CRYAB;HSPD1;RPS19;HSP90B1;NT5E;CYBB;SOD2;MAPK1;PIPA;HMGB1;NDUFA13;GNAI2;HYOU1;C3;RELA;STAT3;AOC3;GSK3B;CSPG4;DCN;TGM2;LTA4H;PTK2
macrophage population	Immunology	Y	0.0007	0.0013	Y	-	N	PLIN2;TF;LGALS3;APOA1;APOE;CES1;CAT;FN1;ACAT1;DPP4;ANXA5;LRP1;CD36;STAT3;DCN;TGM2;ELN
complement activation	Immunology	Y	0.0013	-	Y	1.9E-14	Y	ITGB2;ALB;CLU;COMP;HTRA1;PLG;HRG;A2M;C4A;LTF;C5;TF;APOA1;C6;APOE;HP;XDH;VTN;TIH2;MMP2;APP;TXN;FNI;F2;ENO1;GPI;ANXA2;TPI1;CALR;NID1;P4HB;SERPINB1;GAPDH;BGN;HSPD1;HSP90AB1;PHB;SOD2;HMGB1;SERPINA3;EFEMP1;KRT1;C3;VIM;PRSS1;FMOD;PRELP;DCN;HAPLN1

macrophage chemotaxis	ITGB2;CLU;GC;A2M;C4A;LTF;LGALS3;APOA1;S100A4;APP;TXN;ANXA3;F2;MAPK3;RAC1;FBN1;RPS19;CYBB;MAPK1;HMGB1;ROCK1;ELN	Y	0.0052	Y	9.3E-09	C4A;LGALS3;ITGB2;S100A4;A2M;LTF;CLU;APOA1;GC	Y	-
T cell quantity	PTPRC;ITGB2;ALB;LTF;LGALS3;APOE;APP;HLA-DQB1;F2;CALR;P4HB;HSPD1;BGN;ITGAL;STAT3;RELA;GPRIN1;TRIM28;KRT14	Y	0.0149	N	-	-	Y	-
immunity	-	N	-	Y	4.3E-18	HTRA1;APOB;LTF;FLOT2;HRG;MFGE8;RAC2;SDCBP;PLG;HP;ALB;C4A;CTSD;CP;CS TB;TF;IGFBP7;APOE;FABP5;CLU;PTPRC;LGALS3;LGALS3BP;ITGB2;BAX;LCP1;APOA1;SERPINE2;CTSH;A2M;HPX;S100A4;FTH1;GC	Y	-
innate immune response	-	N	-	Y	3.4E-17	APOB;SSC5D;LTF;HRG;MFGE8;RAC2;PLG;HP;ALB;C4A;CTSD;XDH;CP;CSTB;TF;AP OE;CLU;FGB;PTPRC;LGALS3;ITGB2;BAX;LCP1;APOA1;PNP;C7;CTSH;SCARB2;A2M;ITGB3;GC	Y	-
immune response	-	N	-	Y	3.6E-16	C6;PON3;APOB;C5;LTF;HRG;MFGE8;RAC2;PLG;SDCBP;HP;ALB;C4A;CTSD;XDH;CP;CSTB;TF;ALDH1A2;APOE;FABP5;CLU;CE S2;CORO1A;FGB;PTPRC;LGALS3;ITGB2;LGALS3BP;BAX;LCP1;APOA1;PNP;GUSB;S CARB2;HPX;A2M;S100A4;ITGB3;FTH1;GC	Y	-
cell survival	-	N	-	Y	3.8E-16	CA2;HTRA1;NAPA;LTF;COMP;FLOT2;RAC2;MFGE8;PLG;SDCBP;HP;ALB;C4A;CTSD;COL 1A1;XDH;FLII;CSTB;CTBP1;IGFBP7;TF;AP OE;FABP5;FABP7;CLU;PLIN2;CORO1A;PT PRC;DAB2;CTHRC1;ENO2;LGALS3;ITGB2; BAX;LCP1;APOA1;SERPINE2;ACTN1;PNP; HPRT1;A2M;HPX;S100A4;ITGB3;RPL35A;F TH1;GC	Y	-
cell motility	-	N	-	Y	4.8E-16	GPC6;CA2;HTRA1;GMFG;LTF;MFGE8;RAC2;SDCBP;PLG;TWF1;ALB;CTSD;FLII;CST B;TF;APOE;FABP5;FABP7;CLU;CAPG;CO RO1A;DAB2;PTPRC;CTHRC1;LGALS3;ITG B2;LGALS3BP;LCP1;APOA1;ACTN1;C7;FB LN1;S100A4;ITGB3	Y	-
chemotaxis	-	N	-	Y	1.8E-15	GMFG;APOB;COMP;LTF;FLOT2;RAC2;MF GE8;PLG;SDCBP;HP;ALB;C4A;CTSD;TF;A POE;FABP5;CLU;CORO1A;PTPRC;LGALS 3;ITGB2;LCP1;APOA1;ACTN1;GUSB;HPX; S100A4;A2M;ITGB3;FTH1;GC	Y	-
T-cell response	-	N	-	Y	1.4E-09	PTPRC;LGALS3;ITGB2;APOB;LTF;LCP1;A POA1;HP;ALB;CTSD;PNP;TF;APOE;A2M;IT GB3;CLU;FTH1;GC;CORO1A	Y	-

immune system function	Immunology	N	-	Y	1.2E-08	Y	PTPRC;LGALS3;ITGB2;LTF;RAC2;APOA1;MFGE8;HP;ALB;CTSD;C4A;PNP;CP;IGFBP7;APOE;HPX;FTH1;GC	Y	-
hemostasis antigen processing and presentation bacteriostasis	Immunology	N	-	Y	2.0E-08	Y	DAB2;LGALS3;COMP;HRG;MFGE8;APOA1;PLG;ALB;XDH;FBLN1;A2M;ITGB3	Y	-
antigen binding lymphocyte transformation	Immunology	N	-	Y	8.5E-10	Y	LGALS3;ITGB2;BAX;LTF;LCP1;APOA1;RAC2;PLG;HP;ALB;CTSD;XDH;CTSH;A2M;APOE;PLIN2	Y	-
neutrophil phagocytosis	Immunology	Y	0.00017	N	-	N	ALB;A2M;LTF;TF;XDH;CAT;TXN	N	-
cell recognition	Immunology	Y	0.00104	Y	-	Y	ALB;CLU;BAX;HTRA1;PLG;A2M;LTF;CTSH;TF;APOB;ITGB3;APOE;MMP2;APP;FN1;G6PD;HLA-DQB1;ANXA4;F2;ANXA2;FOLH1;EEA1;ANXA1;RHOA;DES;RPL14;MAPK1;HMGB1;PURA;ACSL1;GALNT2;TLN1;NCAM1;TROVE2;TGM2;PTK2	N	-
lymphocyte apoptosis activated T cell adhesion	Immunology	Y	0.00188	Y	7E-13	Y	PTPRC;ITGB2;ALB;LCP1;LTF;LGALS3;CORO1A;HPX;MFGE8;HP;CP;FN1;ANXA5;MAPK3;ANXA1;LRP1;CD36;GAPDH;MAPK1;HMGB1	N	-
phagocyte activity phagocyte behavior	Immunology	Y	0.00304	Y	6E-11	Y	PTPRC;ITGB2;ALB;PLG;LTF;C5;TF;LGALS3;APOA1;APOB;ITGB3;MFGE8;APOE;FN1;HLA-DQB1;HLA-B;PKM;ANXA2;CALR;MSN;STAT1;ANXA1;CDH13;VCAN;CTNNA1;RPL11;LRP1;CD36;PGM3;ACTR2;DES;PRDX5;ACTR3;STAT3;PTPN11;ITGB1;PCDH7;NCAM1;PRAF2;TGM2;MARCKS;NRP1	N	-
	Immunology	Y	0.0025	N	-	N	PTPRC;BAX;PNP;LTF;LGALS3;MFGE8;XDH;TNFRSF6B;F2;GSN;ANXA2;ANXA5;CAV1;STAT1;RAC1;ANXA1;CD36;HMGB1;STAT3;GNAQ;NCAM1;ELN	N	-
	Immunology	Y	0.00203	N	-	N	PTPRC;ITGB2;LGALS3;ITGB3;MSN	N	-
	Immunology	Y	0.00161	Y	1E-11	Y	PTPRC;ITGB2;ALB;RAC2;GUSB;CAPG;GC;PLG;HRG;A2M;LTF;LGALS3;APOA1;ITGB3;MFGE8;APOE;HP;VTN;TNFRSF6B;CAT;APP;FN1;ANXA3;F2;CLIC1;CALR;ANXA5;NID1;MAPK3;RAC1;SNAP23;EIF5A;ANXA1;RHOA;ITGAV;RCN1;CDC42;LRP1;CD36;PRKCD;PPP2R4;HSPD1;PSMD4;PRDX5;CYBB;MAPK1;PIPA;HMGB1;C3;ROCK1;STAT3;VIM;ROCK2;DAG1;DPYSL3;ARF6;RAP1A;INPP4A;MYLK;PTK2;DYSF;ELN	N	-
	Immunology	Y	0.00304	Y	-	N	ITGB2;CAPG;PLG;LTF;LGALS3;MFGE8;HP;APP;FN1;LRP1;CYBB;HMGB1;ARF6;DYSF	N	-

macrophage recognition	Immunology	Y	0.00374	CLU;APOB;VTN;CAT;FN1;NCL;CALR;ANXA5;CD36;TGM2	N	-	-	N	-
mast cell adhesion	Immunology	Y	0.00375	ALB;NAGA;LGALS3;ITGB3;VTN;F2;ITGAV;ITGB1;FES	N	-	-	N	-
neutrophil retention	Immunology	Y	0.00403	ITGB2;RAC2;XDH;CAT;FN1;MSN;RAC1PTPRC;ITGB2;ALB;RAC2;LCP1;LTF;LGALS3;APOA1;CORO1A;HPX;HP;XDH;VTN;CAT;ITIH2;MMP2;TXNRD1;APP;TXN;FN1;F2;ADK;NID1;MAPK3;CAV1;TGFB1;ANXA1;RHOA;P4HB;CD9;STX4;PRKCD;SO D2;MAPK1;PIIA;HMGB1;NT5E;ITGAL;STAT3;AOC3;ROCK1;TLN1;ITGB1;RAP1A;MYLK;TGM2;MARC KS;PTK2	N	-	-	N	-
neutrophil adhesion	Immunology	Y	0.00495	ITGB2;ALB;LTF;XDH;MMP2;F2;MAPK3;RAC1;MAPK1;PTK2	N	-	-	N	-
neutrophil aggregation	Immunology	Y	0.00505	PTPRC;ITGB2;ALB;BAX;GC;LCP1;A2M;LTF;C5;TF;LGALS3;CORO1A;LGALS3BP;ITGB3;APOE;HP;TNFRSF6B;CAT;MMP2;APP;FN1;HLA-B;F2;ARL8B;DPP4;PRDX1;CALR;MYH9;MAPK3;SE T;STAT1;RAC1;EIF5A;ANXA1;P4HB;CDC42;PGM3;GRB2;PRDX2;PRKCD;CTSC;TALDO1;HSPD1;HSP90B1;MAPK1;HMGB1;CRK;HYOU1;ITGAL;DNM2;DLAT;C3;STAT3;VIM;GSK3B;PARVA;TUBB;PTPN11;KIF5B;NCAM1;HSPB1;ASPH;SVIL;DMD;SERPINH1;CRKL;NRP1;ELN	N	-	-	N	-
NK cell mediated cytotoxicity	Immunology	Y	0.00574	PLIN2;ALB;PLG;ALDH1A2;FTH1;LTF;LGALS3;APOA1;APOE;HP;TNFRSF6B;MMP2;APP;FN1;F2;FKBP1A;DPP4;MAPK3;STAT1;RAC1;ANXA1;RHOA;NDRG2;ITGAV;CDC42;FSCN1;CD36;HSPD1;PHB;MAPK1;HMGB1;STAT3;GSK3B;PTPN11;TGM2;NRP1	N	-	-	N	-
dendritic cell behavior	Immunology	Y	0.0065	ITGB2;HRG;C4A;LGALS3;FN1;CALR;C3	N	-	-	N	-
immune complex clearance	Immunology	Y	0.00671	RAC2;LGALS3;APOA1;APOE;TXN;STAT1;CD36PTPRC;ITGB2;ALB;LTF;LGALS3;CTSD;MFGE8;HSP E1;DPP4;RAC1;RCN1;LRP1;CDC42;HSPD1;STAT3	N	-	-	N	-
macrophage phenotype	Immunology	Y	0.00713	ITGB2;COMP;PLG;C4A;APP;TXN;MPZ;F2;GPI;STAT3;C3;FMOD	N	-	-	N	-
immunostimulation	Immunology	Y	0.0076	ITGB2;ALB;RAC2;LTF;LGALS3;APOA1;FN1;MAPK3;SNAP23;ANXA1;PRKCD;MAPK1	N	-	-	N	-
complement activation, alternative pathway	Immunology	Y	0.00818	ITGB2;ALB;RAC2;LTF;LGALS3;APOA1;FN1;MAPK3;SNAP23;ANXA1;PRKCD;MAPK1	N	-	-	N	-
neutrophil degranulation	Immunology	Y	0.00818	ITGB2;ALB;RAC2;LTF;LGALS3;APOA1;FN1;MAPK3;SNAP23;ANXA1;PRKCD;MAPK1	N	-	-	N	-

neutrophil chemotaxis	Y	0.00828	LN	N	-	-	N	-	-	PTPRC;ITGB2;RAC2;GUSB;GC;C4A;LTF;TF;LGAL S3;HPX;GMFG;HP;VTN;APP;SSR1;TXN;FN1;F2;S OD3;ENO1;PSMA3;NID1;MAPK3;CAV1;MSN;STAT 1;RAC1;YARS;MAP2K1;RHOA;SBD5;CDC42;CAP N2;GAPDH;RPS19;MAPK1;PIPA;HMGB1;SERPINA 3;RELA;ROCK1;GNA13;STAT3;GNB2;GSK3B;GNA Q;RAP1A;HSPB1;SPARC;TGM2;MARCKS;CRKLE
cytolysis	Y	0.00944	LN	N	-	-	N	-	-	PTPRC;ITGB2;ALB;CLU;BAX;LCP1;A2M;C4A;FTH1 ;LTF;TF;LGALS3;C6;CTSD;CP;VTN;CAT;PEPD;PS ME1;APP;G6PD;HLA-B;F2;ARL8B;NQO1;DPP4;CALR;ANXA5;MAPK3;LM NB1;GSR;STAT1;MAP2K1;GARS;RHOA;HSPA5;G RB2;ACTR2;GAPDH;TCEB1;HSPD1;EEF2;MAPK1; HMGB1;SERPINA3;SH3BGR1;ACTR3;ITGAL;C3; KRT5;STAT3;GNB2;ATP2B4;PRELP;ATP5A1;NCA M1;HSPB1;ADD1;KRT10;KRT14
immunocompetent cell adhesion	Y	0.00956		N	-	-	N	-	-	PTPRC;ITGB2;APP;FN1;GSK3B;PTK2
phagocyte recognition	Y	0.01032		N	-	-	N	-	-	ITGB2;LGALS3;MFGE8;VTN;CD36
lymphocyte proliferation	Y	0.0114		N	-	-	N	-	-	PTPRC;ITGB2;ALB;A2M;FTH1;LTF;TF;CORO1A;A POE;HP;CP;TNFRSF6B;CAT;PEPD;APP;FN1;F2;S TT3A;FABP4;DPP4;ANXA2;HBB;CALR;STAT1;ANX A1;CD9;HSPD1;TALDO1;IMPDH2;MAPK1;HSP90B 1;HMGB1;ITGAL;STAT3;RELA;GSK3B;ITGB1;GNA S;NCAM1
blood platelet phagocytosis	Y	0.014		N	-	-	N	-	-	ITGB2;ITGB3;MFGE8;APP;CD36
eosinophil function	Y	0.01435		N	-	-	N	-	-	PTPRC;ITGB2;RAC2;TXN;FN1;MAPK1
rosetting	Y	0.01486		N	-	-	N	-	-	ALB;A2M;CP;FN1;CD36
lymphocyte quantity	Y	0.01532		N	-	-	N	-	-	ITGB2;ALB;TF;APOA1;CP;TNFRSF6B;HSPE1;SOD 3;SERPINB1;ITGAL;AOC3;TES
Th1 Th2 balance	Y	0.01588		N	-	-	N	-	-	A2M;LTF;LGALS3;S100A4;HP;APP;GSN;CALR;PPI A;STAT3
granulocyte function	Y	0.0159		N	-	-	N	-	-	ITGB2;LCP1;LTF;LGALS3;HP;FN1;DBNL;ITGAL;HS PB1;SERPINH1
T cell death	Y	0.01694		N	-	-	N	-	-	PTPRC;BAX;LTF;TF;LGALS3;CTSD;CAT;FN1;SNA P23;AIFM1;MAPK1;STAT3;MAT2B
blast transformation	Y	0.01777		N	-	-	N	-	-	ALB;A2M;FN1;F2;DPP4;SPR

cell phagocytosis	Immunology	N	-	-	4E-10	Y	LGALS3;ITGB2;LTF;HRG;MFG8;RAC2;HP;ALB;C4A;TF;A2M;APOE;ITGB3;CORO1A	N	-
T cell migration	Immunology	N	-	-	3E-09	Y	PTPRC;LGALS3;ITGB2;GMFG;LCP1;RAC2;APOA1;SDCBP;ACTN1;ALB;ITGB3;CORO1A	N	-
cell development	Immunology, Lipid-related	N	-	-	2.2E-13	Y	CA2;HTRA1;LTF;C5;COMP;DDX6;MFG8;RAC2;SDCBP;HP;ALB;XDH;CTBP1;ALDH1A2;TF;APOE;FABP7;CLU;PLIN2;CORO1A;PTPRC;DAB2;LGALS3;ENO2;ITGB2;BAX;APOA1;ACTN1;PNP;FBLN1;S100A4;ITGB3;FTH1	Y	-
endothelial cell adhesion	Immunology, Vascular	N	-	-	4.2E-09	Y	DAB2;LGALS3;ITGB2;LTF;HRG;APOA1;MFG8;PLG;ALB;XDH;CP;APOE;CLU;ITGB3;CA2;GPC6;HTRA1;GMFG;CCDC80;LTF;NAPA;COMP;FLOT2;MFG8;RAC2;PLG;SDCBP;ALB;COL1A1;CTSD;FLI1;XDH;CTBP1;TF;IGFBP7;APOE;FABP5;CLU;FABP7;CAPG;PTPRC;DAB2;CTHRC1;LGALS3;LGALS3B;ITGB2;BAX;LCP1;SERPINE2;ACTN1;PNP;GUSB;CTSH;FBLN1;HPX;A2M;S100A4;ITGB3	Y	Y
cell invasion	Immunology, Vascular	N	-	-	2.5E-17	Y	GPC6;HTRA1;GMFG;CCDC80;COMP;LTF;FLOT2;HRG;RAC2;MFG8;PLG;SDCBP;HP;TF;TW1;ALB;CTSD;COL1A1;FLI1;XDH;CTBP1;CSTB;TF;IGFBP7;APOE;FABP5;FABP7;CLU;CAPG;PTPRC;DAB2;CTHRC1;LGALS3;ENO2;ITGB2;LGALS3BP;LCP1;APOA1;SERPINE2;ACTN1;CTSH;FBLN1;HPRT1;HPX;A2M;S100A4;ITGB3;FTH1	Y	Y
cell migration	Immunology, Vascular	N	-	-	1.0E-13	Y	HTRA1;COMP;LTF;FLOT2;HRG;DDX6;RAC2;SDCBP;PLG;HP;TW1;ALB;CTBP1;IGFBP7;APOE;FABP7;CLU;CAPG;PLIN2;DAB2;PTPRC;CTHRC1;LGALS3;ITGB2;BAX;APOA1;SERPINE2;ACTN1;C7;HPRT1;FBLN1;A2M;S100A4;ITGB3	Y	-
cell phenotype	Immunology, Vascular	N	-	-	1.2E-11	Y	CA2;NAPA;LTF;FLOT2;HRG;GLB1;HP;ALB;COL1A1;C4A;CTSD;XDH;CP;CTBP1;ALDH1A2;IGFBP7;TF;APOE;FABP5;CLU;PLIN2;DAB2;PTPRC;CTHRC1;LGALS3;ITGB2;BAX;BAX;LCP1;APOA1;SERPINE2;FBLN1;A2M;S100A4;HPX;FTH1;GC	Y	-
cell quantity	Immunology, Vascular	N	-	-	2.1E-10	Y		Y	-

cell adhesion umbilical vein	Immunology, Vascular	N	-	3.8E-21	Y	HTRA1;GMFG;CCDC80;COMP;LTF;FLOT2; HRG;RAC2;MFGE8;PLG;SDCBP;HP;ALB;C TSD;C4A;XDH;FLII;CP;FGG;FBLN7;IGFBP 7;TF;APOE;FABP5;CLU;FABP7;CORO1A;D AB2;PTPRC;CTHRC1;LGALS3;LGALS3BP;I TGB2;BAX;LCP1;APOA1;SERPINE2;ACTN 1;GUSB;SCARB2;ITIH1;FBLN1;HPX;A2M;S 100A4;ITGB3	Y	-
endothelial cell adhesion	Immunology, Vascular	Y	0.00338	-	N	CLU;ITGB3;VTN;FN1;F2;TGFB1;CD36;ITGB1;SPAR C;DPT;FBLN5;ELN	N	-
cell contact	Immunology, Vascular	N	-	3E-09	Y	DAB2;PTPRC;LGALS3;ITGB2;LGALS3BP;L TF;LCP1;DECR1;APOA1;PLG;CTSD;S100A 4;CLU;ITGB3	N	-
cell formation	Immunology; Lipid-related	N	-	4.0E-16	Y	GMFG;LTF;C5;MFGE8;RAC2;SDCBP;PLG; ALB;COL1A1;CTSD;XDH;FLII;ALDH1A2;IG FBP7;TF;APOE;FABP5;CLU;PLIN2;CORO1 A;PTPRC;DAB2;CTHRC1;LGALS3;ENO2;R PL36A;ITGB2;BAX;LCP1;APOA1;SERPINE 2;ACTN1;SCARB2;FBLN1;HPRT1;A2M;S10 0A4;ITGB3;FTH1	Y	-
cell interaction	Immunology; Vascular	Y	0.0152	1.4E-15	Y	FABP7;PTPRC;ITGB2;ACTN1;ALB;CLU;BAX;HTRA 1;PLG;HRG;A2M;LTF;TF;LGALS3;FLII;APOA1;SCA RB2;LGALS3BP;APOB;ITGB3;SERPINE2;MFGE8;A POE;HP;DAB2;VTN;CAT;MMP2;APP;TXN;FN1;HLA B;NCL;F2;AQP1;SOD3;FABP4;DPP4;S100A10;ANX A2;CALR;ANXA5;SERPINF1;NID1;MAPK3;CAV1;S TAT1;TGFB1;MSN;RAC1;TPT1;ANXA1;PP1B;CDH1 3;RHOA;ITGAV;CTNNA1;P4HB;VCAN;CD9;CDC42; LRP1;CD36;FSCN1;PRKCD;GAPDH;BGN;CRYAB; EPRS;PAK2;AIP;HSPD1;RECK;ITGA5;FAP;NT5E;M APK1;PP1A;HMGB1;IQGAP1;CYBB;SOD2;VASP;H YOU1;PPP2R1A;ANXA6;CRK;ACSL1;HSPG2;EEF1 A1;ITGAL;STAT3;AOC3;ROCK1;VCL;VIM;DAG1;TL N1;GSK3B;ITGB1;RRAS;COL4A1;ARF6;FUBP1;CS PG4;ITGB5;RAP1A;JUP;MFAP4;COL4A2;DCN;NCA M1;PPP1R12A;SPARC;ILK;CTTN;PPP1R14A;TGM 2;RPL12;ENAH;MRV1;SYNM;MARCKS;PTK2;FBL N5;DYSF;CRKL;NRP1;ELN	Y	-

cell differentiation	Immunology; Vascular; Lipid-related	N	-	0.0007	ALB;APOA1;APOE;CP;CAT;APP	ENO2;ALB;CLU;BAX;PON3;ACOX1;LTF;TF;APOA1;APOB;APOE;HP;CP;CAT;APP;DPP4;DLD;GSR;HS;D17B10;AIFM1;PRDX2;ATIC;PRDX3;SOD2;ARF1;S;TAT3;GPD1;STK25;GNAS;GPX1;ACADM	Y	8.2E-20	GMFG;CCDC80;LTF;COMP;DDX6;MFGE8;PLG;ALB;CP;TF;ALDH1A2;IGFBP7;APOE;FABP5;CLU;DAB2;ETHE1;CTHRC1;LGALS3;APMAP;BAX;TANC1;ACTN1;CTSH;FBLN1;A2M;HPX;S100A4;CA2;PABPC4;HTRA1;APOB;FLOT2;HNMT;RAC2;CLEC3B;SDCBP;HP;CTSD;COL1A1;C4A;XDH;FLI1;CSTB;FBLN7;CTBP1;FABP7;TBC1D10C;CAPG;ACOX1;PLIN2;CORO1A;PTPRC;AOX1;ENO2;LGALS3;HEXA;PON3;BAX;CCDC80;APOB;LTF;MFGE8;APOA1;PLG;ALB;XDH;CTBP1;ALDH1A2;TF;SCARB2;APOE;FABP5;CLU;FABP7;ACOX1;PLIN2	Y	-	Y
lipoprotein oxidation	Lipid-related	Y	-	0.0007	ALB;APOA1;APOE;CP;CAT;APP	ENO2;ALB;CLU;BAX;PON3;ACOX1;LTF;TF;APOA1;APOB;APOE;HP;CP;CAT;APP;DPP4;DLD;GSR;HS;D17B10;AIFM1;PRDX2;ATIC;PRDX3;SOD2;ARF1;S;TAT3;GPD1;STK25;GNAS;GPX1;ACADM	N	-	Y	Y	Y	Y
lipid transport	Lipid-related	N	-	-	-	ENO2;ALB;CLU;BAX;PON3;ACOX1;LTF;TF;APOA1;APOB;APOE;HP;CP;CAT;APP;DPP4;DLD;GSR;HS;D17B10;AIFM1;PRDX2;ATIC;PRDX3;SOD2;ARF1;S;TAT3;GPD1;STK25;GNAS;GPX1;ACADM	Y	3.8E-10	APOB;LTF;APOA1;MFGE8;PLG;ALB;SCARB2;TF;APOE;FABP5;FABP7;CLU;PLIN2	Y	Y	Y
lipid storage	Lipid-related	N	-	-	-	ENO2;ALB;CLU;BAX;PON3;ACOX1;LTF;TF;APOA1;APOB;APOE;HP;CP;CAT;APP;DPP4;DLD;GSR;HS;D17B10;AIFM1;PRDX2;ATIC;PRDX3;SOD2;ARF1;S;TAT3;GPD1;STK25;GNAS;GPX1;ACADM	Y	6.9E-09	LGALS3;HEXA;PON3;BAX;CCDC80;APOB;LTF;MFGE8;APOA1;PLG;ALB;XDH;CTBP1;ALDH1A2;TF;SCARB2;APOE;FABP5;CLU;FABP7;ACOX1;PLIN2	Y	Y	Y
lipid oxidation	Lipid-related	Y	0.0057	-	ENO2;ALB;CLU;BAX;PON3;ACOX1;LTF;TF;APOA1;APOB;APOE;HP;CP;CAT;APP;DPP4;DLD;GSR;HS;D17B10;AIFM1;PRDX2;ATIC;PRDX3;SOD2;ARF1;S;TAT3;GPD1;STK25;GNAS;GPX1;ACADM	ENO2;ALB;CLU;BAX;PON3;ACOX1;LTF;TF;APOA1;APOB;APOE;HP;CP;CAT;APP;DPP4;DLD;GSR;HS;D17B10;AIFM1;PRDX2;ATIC;PRDX3;SOD2;ARF1;S;TAT3;GPD1;STK25;GNAS;GPX1;ACADM	Y	1.7E-09	ENO2;BAX;PON3;APOB;LTF;APOA1;HP;ALB;CP;TF;APOE;CLU;ACOX1	Y	Y	99
lipid metabolism	Lipid-related	N	-	-	ITGB2;PLIN2;PON3;PLG;APOA1;CORO1A;SCARB2;APOB;CTSD;APOE;XDH;CES1;UFM1;CAT;MMP2;APP;AEBP1;FN1;ACAT1;F2;FABP4;DPP4;PRDX1;CALR;MAPK3;CAV1;STAT1;RHOA;LRP1;CD36;HSPA5;GRB2;PRKCD;CAPN1;APPL1;MAPK1;PIPA;APOA1BP;DNM2;RELA;ROCK1;VIM;GSK3B;ROCK2;GPX1;HSPB1;LSS;PLIN1;PXDN	ITGB2;PLIN2;PON3;PLG;APOA1;CORO1A;SCARB2;APOB;CTSD;APOE;XDH;CES1;UFM1;CAT;MMP2;APP;AEBP1;FN1;ACAT1;F2;FABP4;DPP4;PRDX1;CALR;MAPK3;CAV1;STAT1;RHOA;LRP1;CD36;HSPA5;GRB2;PRKCD;CAPN1;APPL1;MAPK1;PIPA;APOA1BP;DNM2;RELA;ROCK1;VIM;GSK3B;ROCK2;GPX1;HSPB1;LSS;PLIN1;PXDN	Y	6.3E-15	ENO2;BAX;PON3;APOB;LTF;APOA1;HP;ALB;CP;TF;APOE;CLU;ACOX1 CA2;HTRA1;PON3;APOB;LTF;DDX6;HRG;PLG;HP;ALB;COL1A1;C4A;CTBP1;TF;APOE;FABP5;CLU;FABP7;CES2;ACOX1;PLIN2;AOX1;ENO2;APMAP;DECR1;APOA1;A2M;ITGB3;GC	Y	Y	Y
foam cell formation	Lipid-related	Y	0.0071	-	ITGB2;PLIN2;PON3;PLG;APOA1;CORO1A;SCARB2;APOB;CTSD;APOE;XDH;CES1;UFM1;CAT;MMP2;APP;AEBP1;FN1;ACAT1;F2;FABP4;DPP4;PRDX1;CALR;MAPK3;CAV1;STAT1;RHOA;LRP1;CD36;HSPA5;GRB2;PRKCD;CAPN1;APPL1;MAPK1;PIPA;APOA1BP;DNM2;RELA;ROCK1;VIM;GSK3B;ROCK2;GPX1;HSPB1;LSS;PLIN1;PXDN	ITGB2;PLIN2;PON3;PLG;APOA1;CORO1A;SCARB2;APOB;CTSD;APOE;XDH;CES1;UFM1;CAT;MMP2;APP;AEBP1;FN1;ACAT1;F2;FABP4;DPP4;PRDX1;CALR;MAPK3;CAV1;STAT1;RHOA;LRP1;CD36;HSPA5;GRB2;PRKCD;CAPN1;APPL1;MAPK1;PIPA;APOA1BP;DNM2;RELA;ROCK1;VIM;GSK3B;ROCK2;GPX1;HSPB1;LSS;PLIN1;PXDN	N	-	APOB;LTF;NAPA;FLOT2;RAC2;MFGE8;SDCBP;PLG;HP;TWf1;ALB;TWf2;CTSD;EHD4;XDH;CTBP1;TF;APOE;TBC1D10C;CLU;DAB2;PTPRC;LGALS3;ITGB2;HEXA;BAX;APOA1;SERPINE2;ACTN1;GUSB;SCARB2;A2M;HPX;ITGB3;GC	Y	Y	Y
non-selective vesicle endocytosis	Lipid-related	N	-	-	-	ITGB2;PLIN2;PON3;PLG;APOA1;CORO1A;SCARB2;APOB;CTSD;APOE;XDH;CES1;UFM1;CAT;MMP2;APP;AEBP1;FN1;ACAT1;F2;FABP4;DPP4;PRDX1;CALR;MAPK3;CAV1;STAT1;RHOA;LRP1;CD36;HSPA5;GRB2;PRKCD;CAPN1;APPL1;MAPK1;PIPA;APOA1BP;DNM2;RELA;ROCK1;VIM;GSK3B;ROCK2;GPX1;HSPB1;LSS;PLIN1;PXDN	Y	3.7E-19	APOB;LTF;NAPA;FLOT2;RAC2;MFGE8;SDCBP;PLG;HP;TWf1;ALB;TWf2;CTSD;EHD4;XDH;CTBP1;TF;APOE;TBC1D10C;CLU;DAB2;PTPRC;LGALS3;ITGB2;HEXA;BAX;APOA1;SERPINE2;ACTN1;GUSB;SCARB2;A2M;HPX;ITGB3;GC	Y	-	-

internalization receptor mediated endocytosis	Lipid-related	N	-	3.6E-14	Y	PABPC4;APOB;LTF;FLOT2;HRG;RAC2;MFG E8;SDCBP;PLG;HP;TWF1;ALB;EHD4;CP;TF;APOE;CLU;DAB2;PTPRC;LGALS3;ITGB 2;NAGA;BAX;APOA1;SERPINE2;ACTN1;G USB;SCARB2;A2M;ITGB3;FTH1	Y	-
opsinization	Lipid-related	N	-	2.0E-10	Y	DAB2;TF;APOB;HPX;LTF;A2M;APOE;CLU; HP;TWF1;ALB	Y	-
cholesterol esterification	Lipid-related	Y	0.00066	6E-09	Y	ITGB2;ALB;PLG;C4A;LGALS3;APOB;MFG E8;APOE ;VTN;APP;FN1;ANXA2;CALR;ITGA5;C3;VIM	N	-
lipid hydrolysis	Lipid-related	Y	0.01035	-	N	ALB;A2M;APOA1;APOB;APOE;HP;CES1;APP;OSB P;ACAT1;CAV1;LRP1;HSPA5;RAB7A;MAPK1;VIM; LSS	N	-
adipocyte differentiation	Lipid-related	Y	0.0155	-	N	PLIN2;ALB;APOA1;APOE;CES1;FN1;F2;MGLL;ARL 8B;FABP4;ANXA5;CD36;MAPK1;PRDX6;ACSL1;G NAQ;ARF6;PLCD1;MARCKS;PLIN1	N	-
adipogenesis	Lipid-related	N	-	2E-12	Y	CTHRC1;APMAP;LGALS3;LGALS3BP;CCD C80;LTF;APOA1;CLEC3B;ALB;CTSD;XDH; FLI1;CP;IGFBP7;TF;A2M;APOE;FABP5;AC OX1;PLIN2;FTH1	N	-
non-selective vesicle exocytosis	Lipid-related	N	-	2E-08	Y	AOX1;CTHRC1;APMAP;LGALS3;HTRA1;B AX;CCDC80;LTF;SDCBP;PLG;ALB;CTSD;X DH;CTSH;TF;APOE;FABP5;ACOX1;PLIN2 DAB2;PTPRC;ITGB2;GMFG;BAX;LTF;NAP A;HNMT;MFG E8;RAC2;APOA1;PLG;ALB;C TSD;C4A;CTBP1;TF;S100A4;APOE;CORO 1A	N	-
ingestion postmenopausal use	Miscellaneous	Y	0.0019	-	N	PTPRC;PABPC4;ITGB2;ALB;CLU;HRG;TF;APOB;IT GB3;MFG E8;APOE;CP;XDH;HPR1;VTN;CAT;PPP 5C;APP;FN1;G6PD;F2;AQP1;DPP4;CALR;ANXA5; NID1;GSR;RAC1;RPLP0;RHOA;ITGAV;CDC42;LRP 1;CD36;RAB7A;ITGA5;CYBB;SOD2;HMGB1;LUM;A DH1B;ALDH2;ROCK1;ITGB1;AGO1;SGCE;ARF6;G PX1;ILK;MYLK;ASPH;FBP1;MARCKS;PTK2	Y	-
	Miscellaneous	Y	0.0023	-	N	APOB;APOE;HP;COL1A1;CAT;MMP2;TXN1;F2;MAP K1;EPHX2;ELN	Y	-

neuronal death	Miscelleneauo	N	-	4.4E-09	Y	APOB;LTF;MFGE8;PLG;CLEC3B;ALB;CTS D;XDH;CP;CSTB;TF;APOE;CLU;LGALS3;E NO2;BAX;SERPINE2;GUSB;CTSH;HPRT1; S100A4;HPX;A2M;GC	Y	-
hemato-encephalic barrier	Miscelleneauo	N	-	7.6E-09	Y	ENO2;LGALS3;APOB;LTF;HNMT;APOA1;P LG;SERPINE2;ACTN1;ALB;XDH;PNP;CP;T F;APOE;S100A4;CLU	Y	-
cell aggregation	Miscelleneauo	N	-	2.3E-08	Y	PTPRC;LGALS3;ITGB2;LGALS3BP;SDCBP ;PLG;HP;ALB;CP;IGFBP7;TF;A2M;CLU;ITG B3;CAPG	Y	-
aging	Miscelleneauo	N	-	6.8E-13	Y	HTRA1;APOB;COMP;RAC2;MFGE8;HP;AL B;CTSD;COL1A1;XDH;CP;FGG;TF;IGFBP7; APOE;FABP5;CLU;ACOX1;PLIN2;PTPRC;C THRC1;LGALS3;BAX;APOA1;GUSB;A2M;H PX	Y	-
bone resorption	Miscelleneauo	N	-	4.4E-09	Y	CA2;PTPRC;CTHRC1;LGALS3;LTF;APOA1 ;RAC2;MFGE8;HP;ALB;CTSD;COL1A1;XD H;GUSB;CSTB;S100A4;ITGB3;GC	Y	-
intestinal absorption	Miscelleneauo	Y	4.9E-05	-	N	ALB;GUSB;TF;APOA1;APOB;APOE;CP;CES2;CES 1;ACAT1;AQP1;ACO1;DPP4;FOLH1;PRMT1;CD36 PLIN2;ALB;A2M;LTF;TF;LGALS3;APOA1;APOB;HP X;APOE;HP;CP;XDH;CAT;APP;FN1;P4HB;LRP1;C D36;AOC3;C3;ITGB5;DCN	N	-
glucose uptake	Miscelleneauo	Y	7.3E-05	-	Y	FTH1;LTF;TF;APOE;CP;XDH;F2;CALR;PRMT1	Y	6E-12
iron absorption	Miscelleneauo	Y	0.00075	-	N	ALB;BAX;GUSB;HTRA1;PLG;IGFBP7;LTF;CTSD;M FGE8;DAB2;PSAP;MMP2;SSR1;AEBP1;F2;SERPIN F1;STAT1;RAC1;RHOA;PRKCD;PHB;MAPK1;CAP N1;STAT3;ITGB1;RAP1A;COL4A2;BIN1;CSRP1;PT K2	N	-
involution	Miscelleneauo	Y	0.00122	-	N	ALB;ACOX1;TF;HP;DAB2;XDH;CAT;AQP1;DPP4;F OLH1;PRKCSH;AHCY;CD36;GNAS;ARF6;ADD1	N	-
renal reabsorption second trimester pregnancy drug transport	Miscelleneauo	Y	0.00224	-	N	ALB;APOA1;COL1A1;FN1;F2;GSN;VCAN;SPR;SER PINH1;ELN	N	-
	Miscelleneauo	Y	0.00237	-	N	ALB;TF;APOE;CAT;MVP;APP;RPS13;CAV1;MSN;R PL23;RHOA;RDX;ELN	N	-
	Miscelleneauo	Y	0.00605	-	N	PLIN2;ALB;CLU;BAX;PLG;IGFBP7;LTF;TF;APOA1; CTSD;MFGE8;APOE;HP;CP;DAB2;XDH;HPRT1;FN 1;G6PD;COX2;AQP1;ACLY;ALAD;ANXA5;HK1;LNP EP;IDH1;CAV1;RAC1;PGM1;HSPA8;YWHAQ;CDC4 2;CD36;GAPDH;ITGA5;PRPS1;UGP2;GLS;SOD2;M APK1;SERPINA3;SELENBP1;STAT3;ACTA2;PTPN 11;DAG1;ITGB1;ACTB;ATP2B4;NCAM1;MYLK;ATP 2A3;FES;PLIN1;PTK2;MYH11	N	-
lactation	Miscelleneauo	Y	0.00631	5E-12	Y	DAB2;BAX;LTF;APOA1;MFGE8;PLG;HP;AL B;CTSD;XDH;CP;IGFBP7;TF;HPRT1;APOE ;CLU;PLIN2	N	-

luteal phase	Miscelleneau	Y	0.00727	PLIN2;BAX;LTF;LGALS3;MFGE8;CP;CAT;MMP2;ANXA4;F2;ITGA1;SOD2;GNB2	N	-	-	N
mammary gland development	Miscelleneau	Y	0.00766	BAX;IGFBP7;S100A4;MFGE8;XDH;MMP2;AEBP1;FN1;RPL6;NIT1;GSN;CAV1;STAT1;PRKCD;PHB2;CRK;STAT3;ITGB1;ATP2B4;RRAS2;ENAH;PTK2;KRIT14;KHDRBS1	N	-	-	N
membrane damage endometrium proliferation	Miscelleneau	Y	0.00861	ITGB2;ALB;CLU;BAX;NAGA;LTF;TF;LGALS3;APOB;CTSD;MFGE8;APOE;XDH;PSAP;CAT;MMP2;APP;S100A11;FN1;G6PD;NQO1;NPM1;HBB;ANXA5;HK1;MAPK3;HBA1;TPT1;ANXA1;PPIB;HSPA8;GPX5;SLC25A5;GAPDH;AIFM1;PRKCD;CAPN1;CYBB;PRDX6;MAPK1;HMGB1;ATP5O;PTRF;GSK3B;TARDBP;ACTA2;DAG1;ATP2B4;GPX1;HSPB1;DMD;DYSF	N	-	-	N
ovulation	Miscelleneau	Y	0.00954	CLU;LTF;LGALS3;NPM1;SERPINF1;CAV1;DCNFABP5;ALB;HTRA1;PLG;A2M;IGFBP7;FLII;APOA1;CTSD;SERPINE2;APOE;XDH;HPRT1;MMP2;FN1;F2;FKBP4;ENO1;PRDX1;DPP4;ANXA2;NID1;MAPK3;CAV1;RHOA;VCAN;CDC42;PGRMC1;PRDX2;RAB7A;GDI1;HSPD1;MAPK1;DCXR;AKR1C1;HSPG2;TLN1;ITGB1;GNAQ;CBR1;HSPB1;DCN;ILK;PDE3A;FERMT2;HAPLN1	N	-	-	N
placenta transfer	Miscelleneau	Y	0.01058	PLIN2;ALB;PLG;TF;CP;MMP2;CD36;ANXA6	N	-	-	N
mucus secretion	Miscelleneau	Y	0.01152	ALB;C4A;LTF;LGALS3;XDH;CAT;MAP2K1;SERPINB1;MAPK1;MARCKS	N	-	-	N
brain microcirculation	Miscelleneau	Y	0.01322	ALB;APOE;APP;F2;GPX1	N	-	-	N
lung permeability	Miscelleneau	Y	0.01351	ITGB2;ALB;HPX;HP;XDH;F2;AQP1;CAV1;TOLLIP;RHOA;CTNNA1;HMGB1;SPARC;MYLK	N	-	-	N
engraftment	Miscelleneau	Y	0.01368	PTPRC;ITGB2;ALB;COMP;RAC2;LCP1;LTF;TF;LGLALS3;ITGB3;FLOT2;APOE;DAB2;VTN;CAT;MMP2;TXN;FN1;HLA-DQB1;F2;DPP4;ANXA2;SERPINF1;CAV1;STAT1;RAC1;GSTP1;RHOA;SBDS;CDC42;PAK2;RECK;ITGA5;CYBB;SOD2;HSP90B1;HMGB1;ITGAL;STAT3;GSK3B;PTPN11;ITGB1;GNAS;IDH2;RAP1A;HSPB1;SPARC;LASP1	Y	3E-09	PTPRC;DAB2;LGALS3;ITGB2;COMP;LTF;FLOT2;LCP1;RAC2;ALB;TF;APOE;ITGB3	N
bacterium adherence	Miscelleneau	Y	0.01645	ALB;COMP;PLG;LTF;TF;FLII;SCARB2;XDH;VTN;CAT;FN1;KRT6A;NCL;F2;ENO1;PGD;GAPDH;HSPD1;SOD2;ITGA1;CRK;VIM;ITGB1;ITGB5;DCN;PTK2	N	-	-	N
fetus growth	Miscelleneau	Y	0.01697	PTPRC;ALB;C5;LGALS3;APOA1;APOE;CP;CAT;TXN;F2;ACLY;DPP4;ALDH9A1;GSR;MAPK3;MTHFD1;LAMA5;IGF2R;MAPK1;NDRG1;GOT2;PDE5A;ACP1;GNAS;ASPH	N	-	-	N

first trimester pregnancy	Miscelleneaou S	Y	0.017	N	-	N	-	ALB;BAX;PLG;LTF;TF;APOA1;HEXA;APOB;ITGB3;SERPINE2;CAT;MMP2;FN1;F2;AQP1;CALR;MAPK3;STAT1;CACYP;GAPDH;YWHAQ;BGN;SOD2;MAPK1;PIIA;VASP;STAT3;C3;RELA;ROCK1;HSPB1;LK;KRT10;TGM2;LTA4H;SERPINH1;FBLN5
sodium reabsorption	Miscelleneaou S	Y	0.01812	N	-	N	-	ALB;LGALS3;HEXA;SERPINE2;PPP5C;AQP1;DPP4;ATP1A1;MAPK3;RAC1;P4HB;ARHGDI2;AHCY;M APK1;CTTN;ADD1
virus attachment	Miscelleneaou S	Y	0.01829	N	-	N	-	PTPRC;ALB;SDCBP;LTF;APOA1;CTBP1;SCARB2;MFGE8;APOE;VTN;APP;FN1;NCL;F2;ILF3;FKBP4;DPP4;ANXA2;ATP1A1;CAV1;RAC1;ANXA1;P4HB;1 TGAV;CD36;HSPA5;GRB2;RAB7A;PIIA;HSP90B1;HSPG2;VIM;DAG1;ARCN1;RPSA;ITGB1;SERPINH1
actin organization	Miscelleneaou S	N	-	Y	4E-15	N	-	GMFG;LTF;FLOT2;RAC2;PLG;SDCBP;HP;TWLF;ALB;TWLF2;CTSD;FLII;XDH;TF;CAPG;CORO1A;PTPRC;CTHRC1;LGALS3;ENO2;ITGB2;LCP1;APOA1;ACTN1;HPX;S100A4;ITGB3;GC
synaptogenesis	Miscelleneaou S	N	-	Y	5E-12	N	-	PTPRC;GPC6;ITGB2;BAX;COMP;LCP1;FL OT2;RAC2;SDCBP;GLB1;SERPINE2;ACTN1;ALB;CTSD;TF;APOE;FABP7;CLU;CORO1A
detoxification (process)	Miscelleneaou S	N	-	Y	6E-11	N	-	ETHE1;AOX1;DHRS4;LTF;HRG;APOA1;HP;ALB;XDH;GUSB;CP;TF;HPX;APOE;ITGB3;CAPG;CES2;FTH1
lumen formation mammary gland involution	Miscelleneaou S	N	-	Y	8E-11	N	-	DAB2;CTHRC1;LGALS3;HTRA1;ITGB2;GMFG;COMP;LTF;HRG;APOA1;PLG;HP;SERPINE2;ALB;COL1A1;CTSD;XDH;FBLN7;ALDH1A2;IGFBP7;S100A4;ITGB3;CLU
developmental process	Miscelleneaou S	N	-	Y	2E-10	N	-	DAB2;CTSD;XDH;BAX;LTF;CLU;MFGE8;PLG
epithelial to mesenchymal transition	Miscelleneaou S	N	-	Y	3E-10	N	-	PTPRC;LGALS3;HTRA1;PABPC4;ITGB2;BAX;COMP;RAC2;PLG;SERPINE2;HP;TWLF1;ACTN1;CTSD;PNP;FLII;TF;ALDH1A2;A2M;APOE;FABP7;ITGB3
morphogenesis	Miscelleneaou S	N	-	Y	2E-09	N	-	HTRA1;GMFG;FLOT2;HRG;MFGE8;PLG;SDCBP;TWLF;ALB;COL1A1;XDH;CTBP1;IGFBP7;CLU;DAB2;CTHRC1;LGALS3;BAX;APOA1;ACTN1;FBLN1;S100A4;ITGB3;FTH1
								DAB2;CA2;CTHRC1;LGALS3;ITGB2;BAX;APOB;PLG;SDCBP;TWLF;ALB;CTSD;CTSH;TF;ALDH1A2;FBLN1;S100A4;APOE;A2M;HPX;CLU;ITGB3

cytoskeleton organization and biogenesis	Miscellaneous	N	-	Y	3E-09	DAB2;CTHRC1;LGALS3;ITGB2;GMFG;LCP1;FLOT2;HRG;RAC2;MFGE8;SDCBP;HP;ALB;ACTN1;FLI1;S100A4;APOE;ITGB3;CAPG;CORO1A	N	-
colony formation	Miscellaneous	N	-	Y	4E-09	CA2;CCDC80;COMP;LTF;RAC2;SDCBP;ALB;CTSD;ALDH1A2;TF;IGFBP7;FABP5;CLU;FABP7;DAB2;CTHRC1;ENO2;LGALS3;RPL36A;BAX;SERPINE2;PNP;FBLN1;S100A4;A2M;ITGB3;FTH1	N	-
viral entry	Miscellaneous	N	-	Y	4E-09	PTPRC;LGALS3;APOB;LTF;APOA1;SDCBP;ALB;CP;SCARB2;APOE;CLU;ITGB3;PLIN2;GC	N	-
viral reproduction	Miscellaneous	N	-	Y	4E-09	APOB;LTF;DDX6;MFGE8;PLG;HP;ALB;CP;CTBP1;CSTB;APOE;CLU;PLIN2;PTPRC;ENO2;LGALS3;ITGB2;LGALS3BP;NMT1;BAX;APOA1;ACTN1;SCARB2;FTH1	N	-
cell localization	Miscellaneous	N	-	Y	5E-09	PTPRC;DAB2;LGALS3;NMT1;BAX;NAPA;LCP1;FLOT2;DDX6;RAC2;SDCBP;TWF1;ACTN1;ALB;SCARB2;TF;HPX;APOE;ITGB3;CAPG;FTH1;CORO1A	N	-
virulence	Miscellaneous	N	-	Y	5E-09	AOX1;HTRA1;ITGB2;HEXA;NMT1;BAX;APOB;LTF;PLG;CLEC3B;ALB;XDH;CP;TF;HPRT1;A2M;CLU	N	-
DNA recombination	Miscellaneous	N	-	Y	9E-09	PTPRC;CA2;BAX;APOB;NAPA;LTF;DDX6;HP;ALB;COL1A1;CTSD;GUSB;CTBP1;HPR1;S100A4;FABP7;ITGB3;ACOX1;CORO1A	N	-
liver development	Miscellaneous	N	-	Y	2E-08	LGALS3;ITGB2;APOB;C5;HP;ALB;IGFBP7;ALDH1A2;A2M;S100A4;ACOX1;PLIN2	N	-
superoxide anion generation	Stress	N	-	Y	4.5E-10	PTPRC;LGALS3;ITGB2;DHRS4;BAX;PON3;LTF;FLOT2;RAC2;APOA1;HP;ALB;COL1A1;CTSD;XDH;GUSB;CP;TF;APOE;A2M;ITGB3;ACOX1;FTH1	Y	Y
cell damage hemagglutination	Stress	N	-	Y	1.9E-08	CA2;APOB;C5;LTF;MFGE8;PLG;HP;ALB;CTSD;XDH;CP;CSTB;TF;APOE;CLU;AOX1;LGALS3;ENO2;ITGB2;BAX;APOA1;ACTN1;PNP;GUSB;A2M;HPX;GC	Y	Y
wound healing	Vascular	Y	0.00022	N	-	COMP;LTF;HRG;MFGE8;RAC2;PLG;CLEC3B;HP;ALB;COL1A1;CTSD;XDH;FLI1;IGFBP7;FABP5;CLU;CAPG;PTPRC;CTHRC1;LGALS3;ITGB2;BAX;FBLN1;A2M;S100A4;ITGB3	N	-
	Vascular	N	-	Y	3.7E-11		Y	Y

vascular endothelium function	Vascular	Y	0.0046	ALB;RAC2;LTF;HP;AQP1;DPP4;TKT;PDE5A;CTTN;HSPA6	N	-	-	CTHRC1;LGALS3;MFGE8;PLG;CLEC3B;SERPINE2;HP;CTSD;CTSH;FBLN1;S100A4;HPX;A2M;CLU	Y	Y
tissue remodeling	Vascular	N	-	-	Y	1.1E-10	-	HTRA1;LTF;COMP;FLOT2;PLG;CLEC3B;HP;ALB;CTSD;XDH;CSTB;APOE;CLU;LGAL S3;BAX;LCP1;APOA1;SERPINE2;ACTN1;GUSB;CTSH;SCARB2;FBLN1;HPX;S100A4;A2M;ITGB3	Y	Y
proteolysis	Vascular	N	-	ITGB2;ALB;CLEC3B;COMP;GUSB;HTRA1;GC;PLG;IGFBP7;FTH1;LTF;TF;LGALS3;APOA1;CTHRC1;ITGB3;S100A4;APOE;FBLN1;COL1A1;CP;XDH;VTN; MMP2;TXN;FN1;FLNB;F2;FKBP1A;GLG1;ANXA2;A LAD;CALR;ANXA5;SERPINF1;MAPK3;PPID;CAV1; STAT1;TGFB1;RAC1;DCLK1;ARL6IP5;YWHA B;COL15A1;RHOA;PGK1;VCAN;CD9;SBDS;CDC42;VPS3 5;LRP1;HSPA9;PRKCD;FBN1;LMNA;BGN;HSPD1; TGA5;NT5E;GLS;MAPK1;SOD2;NPNT;PPP3CA;AL DH7A1;DNM2;HSPG2;ASPN;DSTN;STAT3;AOC3;H MGB2;C3;ROCK1;DPYSL3;GSK3B;PRKAR1A;ACT A2;ITGB1;GNAS;PCDH7;FMOD;PRELP;RAP1A;PD LIM7;HSPB1;ATP2A2;TAGLN;DCN;FERMT2;CHKB; NCAM1;SPARC;COL1A2;CNN1;TGM2;OGN;COL6 A1;ZFYX;HAPLN1;PTK2;NRP1;ELN	Y	5.2E-11	-	Y	Y	
osteogenesis	Vascular	Y	0.0157	-	Y	7.3E-09	-	DAB2;LGALS3;LTF;PLG;HP;CTSD;FLI1;IGF BP7;FBLN1;APOE;S100A4;ITGB3	Y	Y
fibroblast migration	Vascular	N	-	ALB;CLEC3B;PLG;HRG;A2M;APOA1;APOB;FGG;S ERPINE2;VTN;MMP2;APP;FN1;F2;ENO1;S100A10; ANXA2;SERPINF1;ANXA5;LNPEP;CAV1;HBA1;LR P1;IGF2R;GAPDH;FAP;RNPEP;PIIA;HMGB1;SER PINA3;VIM;C3;RAP1A;SPARC;DCN;SERBP1;FBLN 5;ELN	N	-	-	HTRA1;COMP;APOA1;PLG;SERPINE2;HP; CTSD;GUSB;CTSH;A2M;S100A4;HPX;CLU	Y	Y
fibrinolysis	Vascular	Y	0.0038	-	N	-	-	-	Y	Y
ECM degradation	Vascular	N	-	-	Y	4.5E-10	-	-	Y	Y

cell proliferation	Vascular	N	-	-	6.7E-13	Y	CCDC80;COMP;LTF;HRG;DDX6;MFG8;PLG;ALB;CP;IGFBP7;TF;ALDH1A2;APOE;FA;BP5;CLU;ATP6V1E1;DAB2;CTHRC1;LGAL;S3;BAX;LCPI;ACTN1;GUSB;CTSH;FBLN1;S100A4;HPX;A2M;CA2;HTRA1;PON3;APOB;NAPA;FLOT2;HNMT;RAC2;SDCBP;HP;COL1A1;CTSD;XDH;FLI1;CTBP1;CSTB;FABP7;CAPG;CES2;ACOX1;PLIN2;CORO1A;PTPRC;ENO2;RPL36A;ITGB2;LGALS3BP;HEXA;NMT1;DECR1;APOA1;SERPINE2;PNP;HPR1;ITGB3;RPL35A;FTH1;GC	Y	Y
cell growth	Vascular	N	-	-	7.2E-11	Y	CCDC80;LTF;DDX6;HRG;MFG8;PLG;ALB;CP;TF;ALDH1A2;IGFBP7;APOE;FABP5;CLU;DAB2;CTHRC1;LGALS3;BAX;GUSB;FBLN1;S100A4;A2M;HPX;CA2;HTRA1;NAPA;FLOT2;RAC2;SDCBP;HP;AARS;COL1A1;CTSD;XDH;FLI1;CTBP1;FGG;FABP7;PLIN2;PTPRC;ENO2;ITGB2;LGALS3BP;NMT1;APOA1;SERPINE2;PNP;ITGB3;FTH1	Y	Y
capillary permeability	Vascular	Y	0.0183	-	-	N	ITGB2;ALB;COMP;HNMT;PLG;TF;APOB;ITGB3;HPX;XDH;VTN;CAT;MMP2;APP;FN1;F2;PLXDC2;SERPINF1;MAPK3;CAV1;COL15A1;ARHGDI1;RHOA;PRKCD;RDX;SOD2;PTGIS;MAPK1;HMGB1;HSPG2;STAT3;ROCK1;ROCK2;PDE5A;ITGB1;RAP1A;HSPB1;MYLK;DTNA;PTK2;ELN	Y	Y
blood clotting	Vascular	N	-	-	3.1E-12	Y	DAB2;LGALS3;ITGB2;APOB;LTF;HRG;APOA1;MFG8;PLG;HP;SERPINE2;ALB;CTSD;C4A;XDH;FGG;TF;A2M;HPX;ITGB3;GC;FG	Y	Y
blood clot lysis	Vascular	Y	0.0078	-	-	N	ALB;RAC2;PLG;A2M;FGG;SERPINE2;APOE;VTN;MMP2;F2;ANXA2;LRP1;PRKCD;HMGB1;C3;PDE5A	Y	Y
Angiogenesis	Vascular	N	-	-	7.3E-18	Y	CA2;HTRA1;GMFG;APOB;LTF;COMP;FLOT2;HRG;RAC2;MFG8;SDCBP;PLG;HP;TW1;ALB;COL1A1;CTSD;XDH;CP;CTBP1;FGG;FBLN7;TF;ALDH1A2;IGFBP7;APOE;CLU;FGF;DAB2;PTPRC;CTHRC1;LGALS3;LGAL;S3BP;ITGB2;BAX;APOA1;SERPINE2;GUSB;CTSH;FBLN1;ITIH1;HPX;A2M;S100A4;ITGB3;GC	Y	Y
blood circulation	Vascular	Y	0.0002	-	-	N	ALB;PON3;PLG;HRG;TF;APOA1;CTSD;APOE;HP;CP;CAT;MMP2;APP;F2;PKM;GLG1;DPP4;GSTA5;SERPINF1;ANXA5;LNPEP;ANXA1;DES;NDUFA13;HSPG2;C3;KRT5;PDE5A;ILK;KRT14;ELN	Y	-

endothelial cell quantity	Vascular	Y	0.0024	ALB;COMP;APOA1;APOE;APP;ANXA3;F2;DPP4;SERPINF1;ACTB;GPX1;ILK;TGM2	N	-	-	Y	-	LTF;COMP;MFG8;PLG;ALB;CP;IGFBP7;TF;APOE;CLU;DAB2;LGALS3;BAX;LCP1;ACTN1;GUSB;CTSH;FBLN1;S100A4;A2M;HPX;CA2;C6;HTRA1;PON3;APOB;NAPA;FLOT2;RAC2;HP;AARS;COL1A1;CTSD;XDH;CSTB;PLIN2;CORO1A;PTPRC;AOX1;ENO2;ITGB2;NMT1;APOA1;SERPINE2;PNP;C7;HPR1;ITGB3;RPL35A;FTH1;GC
cell death extracellular matrix polymerization	Vascular	N	-	-	Y	1.0E-11	-	Y	9.5E-09	LGALS3;HTRA1;CCDC80;COMP;CLEC3B;SERPINE2;ALB;COL1A1;IGFBP7;FBLN1;APOE;S100A4
blood vessel development	Vascular	N	-	PTPRC;ENO2;ITGB2;ALB;CLU;BAX;RAC2;PLG;A2M;C4A;LTF;TF;APOA1;APOB;HPX;HP;CP;XDH;HPR1;VTN;CAT;APP;TXN;FN1;G6PD;F2;AQP1;LDH A;GPI;TP1;CALR;HBB;ANXA5;RPL19;CAV1;GSR;HBA1;RAC1;ITPA;HSPA9;PRDX2;PHB;HSP90B1;CAPN1;EEF1A1;C3;ROCK1;HADHA;ACPI;PRELPD;CN	Y	8.2E-09	-	Y	2E-10	PTPRC;ENO2;ITGB2;BAX;APOB;LTF;APOA1;RAC2;PLG;HP;ALB;C4A;XDH;CP;TF;HPR1;A2M;HPX;CLU
hemolysis	Vascular	Y	7.6E-06	ACTN1;ALB;CLU;PLG;A2M;LTF;TF;LGALS3;FLOT2;HP;CP;APP;FN1;F2;FABP9;HSPA8;CD36;HSPD1	Y	4E-10	-	Y	-	LGALS3;CP;TF;LTF;A2M;FLOT2;CLU;PLG;HP;ALB;ACTN1
agglutination erythrocyte aggregation fibrin formation osteoclast adhesion	Vascular	Y	2.8E-05	ALB;CLU;A2M;TF;HP;CP;FN1;F2;ALB;HRG;FGG;MFG8;MMP2;FN1;F2;ENO1;ANXA5;P4HB;STAT3;SPARC;DPT	N	-	-	N	-	-
endothelial cell differentiation choroid blood flow vasculature development	Vascular	Y	0.00248	LCP1;ITGB3;VTN;MRV11;PTK2;FBLN7;ALB;LGALS3;ITGB3;HP;DAB2;VTN;TNFRSF8;FN1;F2;PRDX1;DPPP4;MAPK3;CAV1;RAC1;CDH13;RHOA;CDC42;CAPN2;MAPK1;STAT3;VIM;GSK3B;AKAP12;RAP1A;DCN;TAGLN;DDAH2;TGM2;E;DF1;PTK2;NRP1	N	-	-	N	-	-
	Vascular	Y	0.00255	HTRA1;APOE;MMP2;SERPINF1;PDE5A;RAP1A;ELN	N	-	-	N	-	-
	Vascular	Y	0.01399	FBLN1;FN1;SERPINF1;NID1;CD36;LRP1;C14orf16	N	-	-	N	-	-
	Vascular	Y	0.01688	6;PTK2;PALLD;NRP1	N	-	-	N	-	-
cell spreading	Vascular	N	-	PTPRC;DAB2;LGALS3;ITGB2;LGALS3BP;LTF;FLOT2;RAC2;APOA1;PLG;SDCBP;SERPINE2;ACTN1;ALB;FHLI;IGFBP7;FBLN1;A2M;ITGB3	Y	2E-11	-	Y	-	PTPRC;DAB2;LGALS3;ITGB2;LGALS3BP;LTF;FLOT2;RAC2;APOA1;PLG;SDCBP;SERPINE2;ACTN1;ALB;FHLI;IGFBP7;FBLN1;A2M;ITGB3

platelet activation	Vascular, Immunology	N	-	-	1.1E-11	Y	Y	DAB2;PTPRC;CA2;LGALS3;ITGB2;APOB;C5;HRG;APOA1;PLG;SERPINE2;ACTN1;ALB;FGG;TF;SCARB2;APOE;ITGB3;CLU	Y	Y
apoptosis	Vascular, Immunology	N	-	-	9.0E-14	Y	Y	CCDC80;COMP;LTF;C5;HRG;DDX6;MFG8;PLG;ALB;CP;TF;IGFBP7;ALDH1A2;APOE;FABP5;CLU;ETHE1;DAB2;CTHRC1;LGALS3;BAX;ACTN1;GUSB;CTSH;FBLN1;HPX;A2M;S100A4;CA2;HTRA1;PON3;APOB;NAPA;FLOT2;RAC2;SDCBP;CLEC3B;GLB1;HP;C4A;COL1A1;CTSD;XDH;FLI1;CTBP1;CSTB;FGG;CAPG;CES2;ACOX1;PLIN2;CORO1A;PTPRC;ENO2;LGALS3BP;ITGB2;RPL36A;NMT1;APOA1;SERPINE2;PNP;SCARB2;ITGB3;RPL35A;FTH1;GC	Y	Y
cell killing	Vascular, Immunology	Y	0.01553	P1	-	N	N	ITGB2;ALB;CLU;BAX;GUSB;PNP;TF;LGALS3;LGALS3BP;MFG8;XDH;CSTB;VTN;CES1;TNFRSF6B;CAT;MMP2;APP;TXN1;FN1;F2;NQO1;NPM1;SERPINF1;CALR;APEX1;YBX1;MAPK3;SET;GSR;PPID;VDA;C2;STAT1;TPT1;MAP2K1;PI3K;GSTP1;CDC42;IGF2R;HSPA9;HSPA5;AIFM1;PLS3;RECK;EEF2;SLC25A4;SOD2;MAPK1;PPIA;SERPINA3;AK3;ITGAL;C3;STAT3;PDE5A;HDLBP;PRSS1;FMOD;ARF6;CBR1;DCN;NCAM1;HSPB1;SPARC;BIN1;TGM2;PTK2;NR	N	-

Table S1. Supplementary data of Figure 3A.

II Découverte de la cible

II.1 Introduction

Aujourd'hui, il existe un réel besoin de découvrir de nouveaux biomarqueurs afin de développer les outils de diagnostic ainsi que la thérapie de l'athérosclérose. Dans les premiers stades de l'évolution de la plaque, plusieurs types cellulaires sont recrutés parmi lesquels se trouvent les monocytes et macrophages. Ainsi, les monocytes vont coloniser les sites lésionnels par diapédèse après attraction grâce à des chémokines et interaction avec les récepteurs exprimés à la surface des cellules endothéliales. Dans la partie sous-endothéliale, les monocytes sont activés puis se différencient en macrophages. Les macrophages activés expriment des récepteurs scavenger qui participent à la fixation et phagocytose des LDL. Les cellules macrophagiques gorgées de lipides sont alors dites « cellules spumeuses ». Les macrophages ainsi gorgés de lipides entrent en apoptose et sécrètent des molécules fragilisant la plaque telles que les métalloprotéinases. Ces cellules sont donc des cellules clés qu'il est intéressant de cibler grâce à des anticorps humains sélectionnés *in vivo* chez un animal athéromateux. Parmi les cibles du macrophage dans la plaque d'athérome se trouvent de nouvelles molécules considérées comme de nouveaux biomarqueurs potentiels des plaques notamment dite « vulnérables » : les galectines et plus particulièrement la galectine 3^{24,25,26,27}.

Les galectines appartiennent à la famille des bêta-galactoside lectines. Parmi ces galectines, la galectine 3 est impliquée dans beaucoup de pathologies en particulier les maladies chroniques du rein²⁸, le cancer colorectal²⁹, le cancer de la thyroïde³⁰ ainsi que les infections virales^{31,32}. Le taux de galectine 3 est utilisé comme marqueur de risque de maladie chronique du rein²⁸ mais également comme marqueur exprimé dans le cancer colorectal²⁹. La galectine 3 a été utilisée pour le développement d'outils de diagnostic, notamment pour une imagerie PET des tumeurs thyroïdiennes³⁰. Dans le cadre des maladies cardiovasculaires, le rôle de la galectine 3 a été très étudié dans la fibrillation artérielle ou encore dans la cardiomyopathie hypertrophique^{33,34}. Dans l'athérosclérose, la galectine 3 possède un rôle crucial dans l'inflammation notamment par son lien avec le recrutement des monocytes et la phagocytose des LDL par les macrophages. En effet, Zhu *et al.* ont montré que les LDL oxydés et les LDL acétylés marqués à l'iode 125 étaient activement endocytés par les cellules CHO exprimant galectine-3 et que l'incubation avec des LDL acétylés conduisait à une accumulation intracellulaire d'esters de cholestérol, indiquant le rôle de la galectine-3 dans l'endocytose des protéines AGE (Advanced glycation end-product) et des LDL modifiés³⁵. La galectine 3 est également impliquée dans la macrocalcification³⁶. De plus, la galectine 3 possède de nombreux ligands tels que les MMPs³⁷, la galectine 3BP³⁸, le collagène³⁹, les intégrines³⁹, la laminine⁴⁰ et la fibronectine⁴¹ qui sont des protéines surexprimées au niveaux de la plaque.

II.2 Validation *in vitro* d'anticorps humains après reformatage sous format scFv-Fc

Suite à la sélection *in vivo* et au criblage *in vitro* en cytométrie de flux à moyen débit sur des protéines athéromateuses de lapin, les anticorps ont été reformatés et produits dans les cellules mammifères HEK293. La production dans ce système a permis de générer des quantités d'anticorps à une échelle suffisante pour une validation *in vitro* et une identification de la cible. Les anticorps reformatés ont été testés *in vitro* par cytométrie de flux, par immunohistochimie et utilisés dans des expériences d'immunoprécipitation suivie d'une étude en spectrométrie de masse pour déterminer la cible.

II.2.1 Test de réactivité

II.2.1.i La cytométrie de flux

Afin de valider la réactivité des clones scFv-Fc reformatés, la méthode de cytométrie développée pour le criblage des phage-scFv contre des protéines athéromateuses de lapin a été utilisée.

II.2.1.ii Immunohistochimie sur coupes de tissus athéromateux

Après validation de la réactivité par cytométrie de flux, les anticorps sous format scFv-Fc ont été testés sur coupes de tissus athéromateux. Deux modèles précliniques et des biopsies humaines ont été utilisés. Ainsi, les anticorps ont été testés sur des sections d'aorte de lapin de manière à valider les résultats obtenus par cytométrie de flux utilisant des protéines athéromateuses de lapin. Le modèle souris ApoE KO nous a permis de valider l'interaction de l'anticorps afin de l'utiliser ensuite en expérimentation *ex vivo* sur ces mêmes souris pour une imagerie de la plaque. Enfin, les biopsies humaines, obtenues grâce à une collaboration avec le Professeur DUCASSE, chirurgien vasculaire au CHU (Centre hospitalier universitaire) Pellegrin à Bordeaux, nous ont permis de tester la capacité de ces anticorps à reconnaître leur cible sur tissus humains pour une transposition de ces anticorps vers les études cliniques.

II.2.1.iii Immunoprécipitation suivie d'une analyse en spectrométrie de masse

Parmi les anticorps ayant montré les meilleurs résultats en cytométrie de flux et en immunohistochimie sur les deux modèles précliniques lapin et souris et les biopsies humaines, un des anticorps a été sélectionné pour la détermination de la cible. Ainsi, l'immunoprécipitation a été utilisée, notamment par la méthode de « capture directe ». Cette méthode consiste à immunoprécipiter une protéine dans un extrait grâce à l'anticorps d'intérêt préfixé sur un support solide de type billes magnétiques ou billes d'agarose. Afin de capturer la cible, plusieurs étapes sont nécessaires. En premier il s'agit de fixer les anticorps sur les billes magnétiques adsorbées de protéines G. Ensuite le complexe bille/Anticorps est incubé avec des protéines athéromateuses afin d'aller capturer la cible d'intérêt. Le complexe bille/anticorps/protéine cible est passé et retenu sur une colonne fixée sur un aimant. Après de multiples lavages, le complexe est élué par soustraction de l'aimant et étudié en spectrométrie de masse (**Figure 51**).

La spectrométrie de masse permet d'analyser et d'identifier des molécules en fonction de leur masse ainsi que de caractériser leur structure chimique. Le principe de cette méthode repose sur la séparation en phase gazeuse de molécules en fonction de leur rapport masse/charge. Suite à l'introduction des échantillons dans l'injecteur, les molécules sont ionisées par la source d'ionisation. Les ions sont ensuite séparés suivant leur rapport masse/charge par l'analyseur. S'ensuit un détecteur qui compte les ions et amplifie le signal puis un système informatique permettant l'analyse du signal (**Figure 52**).

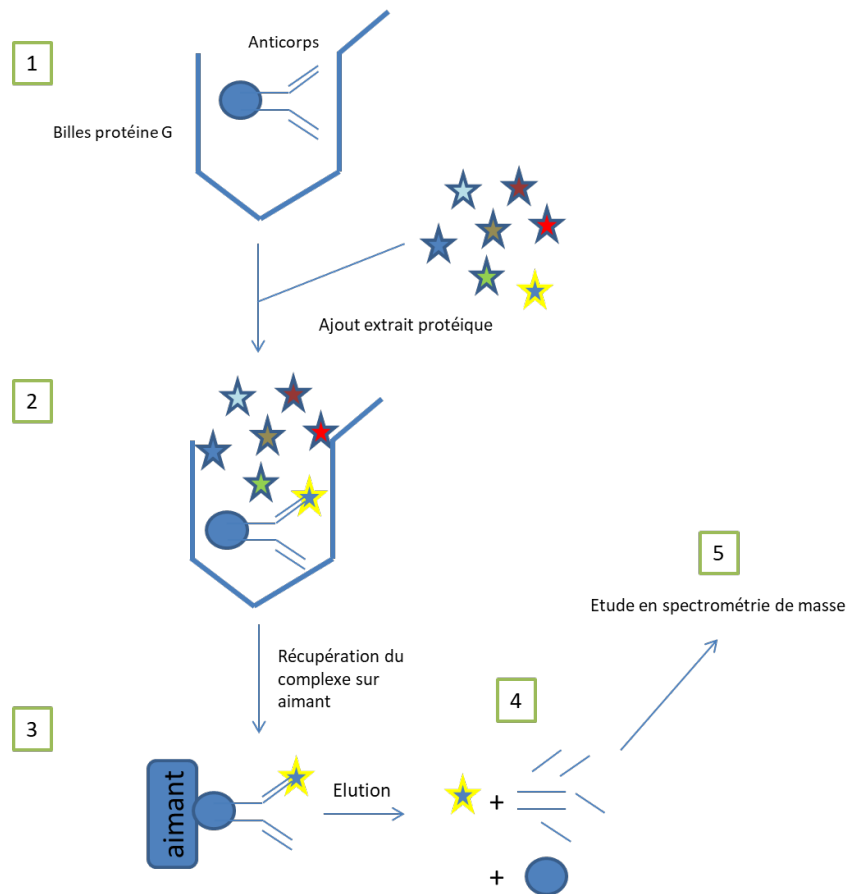


Figure 51 Principe de la méthode d'immunoprécipitation. La capture de l'antigène s'est fait ici par « capture directe ». L'anticorps fixé sur un support solide est mis en présence d'un extrait protéique athéromateux. Après capture, le complexe bille/anticorps/protéine est retenu sur une colonne aimantée puis l'antigène est élué.

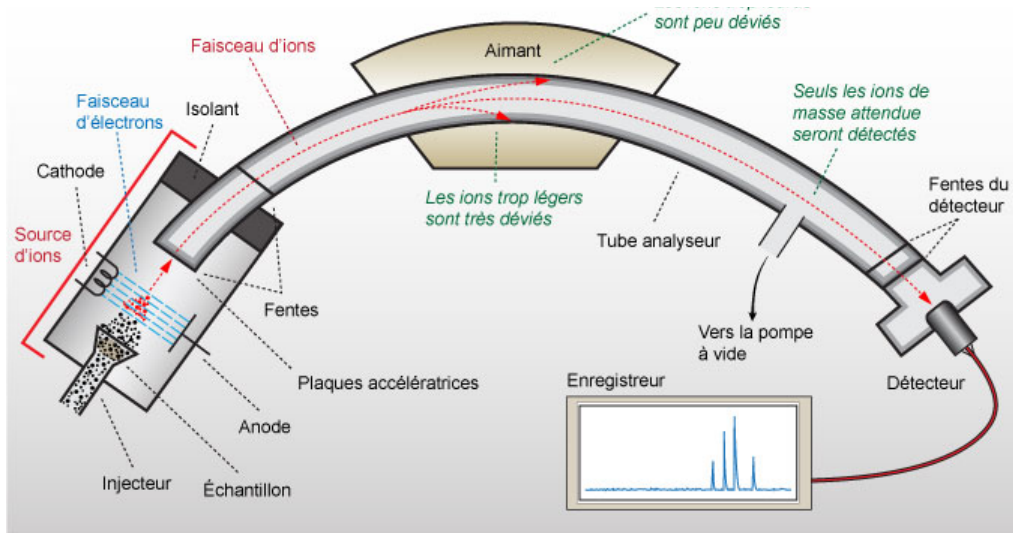


Figure 52 Exemple schématique d'un spectromètre de masse. Le principe de cette méthode repose sur la séparation en phase gazeuse de molécules en fonction de leur rapport masse/charge. Le spectromètre de masse comporte plusieurs composants successifs. En premier se trouve la source d'ionisation suivi par un analyseur séparant les molécules suivant leur rapport masse/charge. S'ensuit un détecteur qui compte les ions et les amplifie puis un système informatique permettant l'analyse du signal (Extrait du site http://www.cima.ualg.pt/piloto/UVED_Geochemie/UVED/site/html/1/1-3/1-3-2/1-3-2-8.html)

II.3 Reformatage d'anticorps en vue d'un couplage aux nanoparticules sous format scFv-Fc-2C

Parmi les anticorps ayant la plus forte spécificité sur les trois modèles, 3 anticorps ont été sélectionnés pour être produits dans les cellules d'insectes. Ce modèle nous a permis d'obtenir

des anticorps au format scFv-Fc avec 2 cystéines supplémentaires nécessaires au couplage à des nanoparticules. Ainsi produits, l'un des trois anticorps, nommé P3, a ensuite permis d'effectuer des tests de colocalisation avec un anticorps commercial dirigé contre la protéine cible identifiée par immunoprécipitation/spectrométrie de masse. L'anticorps P3 couplé à un fluorochrome a également été testé lors d'études préliminaires d'injections *ex vivo* par le cœur de souris euthanasiées en vue de visualiser la plaque d'athérome grâce à l'utilisation d'un microscope fluorescent.

II.3.1.i Immunofluorescence sur coupes de tissus humains

L'immunofluorescence repose sur le même principe que l'immunohistochimie. L'anticorps spécifique d'une protéine présente dans le tissu va se fixer puis être révélé par l'ajout d'un substrat ou grâce à un anticorps fluorescence, ce qui est le cas ici (**Figure 53**). En immunofluorescence, deux anticorps couplés à deux fluorochromes différents peuvent être observés de manière simultanée sur une même lame. Ceci permet donc d'analyser des colocalisations sur tissus. Ainsi, l'anticorps P3 sélectionné *in vivo* puis *in vitro* a été testé en immunofluorescence afin de valider la cible. Pour cela, un anticorps commercial dirigé contre cette même cible a été utilisé.

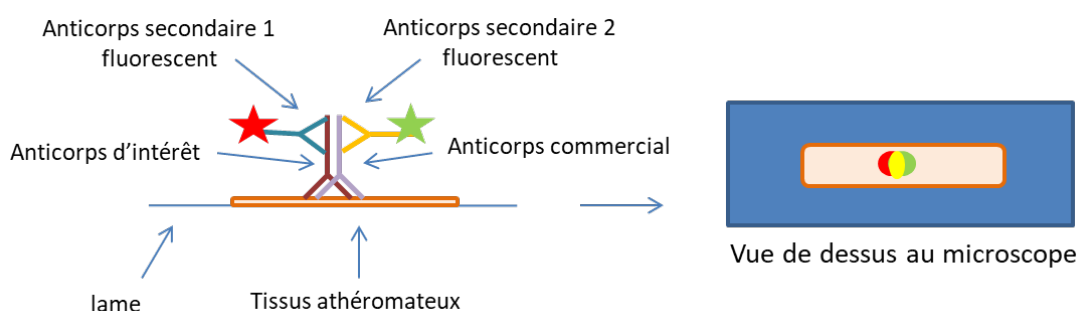


Figure 53 Schématisation du principe d'immunofluorescence. Un tissu préalablement fixé sur une lame est incubé avec les anticorps d'intérêt marqués par des fluorochromes n'émettant pas à la même longueur d'onde. Après de multiple lavage, les lames montées en milieu Vectashield peuvent être observées au microscope fluorescent. Ainsi, la superposition de deux fluorochromes différents permet d'observer des phénomènes de colocalisations sur un tissu.

II.3.1.ii Macrofluorescence sur aortes de souris ApoE KO

Dans l'optique d'utiliser les anticorps sélectionnés à des fins de diagnostic de la plaque d'athérome, l'anticorps P3, dont la cible a été identifiée, a été injecté, après marquage au fluorochrome 647 (rouge lointain), dans des souris ApoE KO. Un anticorps humain au format IgG a été utilisé en tant que contrôle. Après dissection, les aortes ont été incluses dans des gels d'agarose puis observées au microscope fluorescent Nikon.

II.4 Résultats et discussion

L'ensemble des résultats est décrit de manière détaillée dans l'article intitulé « Galectin 3 as a new biomarker of atheroma plaque instability identified by *in vivo* phage display » en cours d'écriture pour être soumis à la revue *Atherosclerosis* et dont je suis l'auteur principal.

Le travail exposé dans cet article décrit une méthode de validation de bio-réactivité d'un anticorps, le clone P3, mais également la découverte de la cible de ce clone, la galectine 3. Il a été démontré ici que l'anticorps P3 reconnaissait une cible présente dans les lésions athéromateuses. Les phage-scFv sélectionnés lors du criblage *in vivo* puis *in vitro*, ayant fait l'objet d'une première publication, ont été reformatés sous format scFv-Fc ou scFv-Fc-2C

dans deux systèmes de production : le système utilisant les cellules mammifères et le système utilisant des cellules d'insectes, respectivement. L'un d'eux, le clone P3, a fait l'objet d'une caractérisation plus poussée. Ainsi, nous avons identifié la cible de P3, la galectine 3, parmi l'extrait de protéines athéromateuses de lapin et d'Homme. L'interaction de P3 avec la galectine 3 a ensuite été validée par immunofluorescence sur coupes de biopsies humaines par comparaison à un anticorps contrôle dirigé contre cette même protéine. La capacité de P3 à interagir avec sa cible sur tissus humains montre qu'il s'agit ici d'une protéine conservée entre le lapin et l'homme. Afin de développer les outils ciblant de diagnostic de la plaque d'athérome par IRM, nous avons pu montrer que cet anticorps P3, après injection *ex vivo* dans la circulation sanguine de souris ApoE KO, était capable de reconnaître sa cible et de marquer la plaque d'athérome, présageant ainsi de son comportement *in vivo* dans le modèle pré-clinique souris.

De plus, la découverte de la cible ici donne accès à de nouvelles voies de traitement non seulement dans l'athérosclérose mais également dans d'autres pathologies comme le cancer. En effet, la galectine 3 identifiée comme nouveau biomarqueur de l'athérosclérose dans des études récentes, a été initialement décrite dans le cancer. Dans les cancers de l'œsophage et de l'estomac, la protéine galectine 3 a été montrée pour être impliquée dans la motilité cellulaire avec un rôle prépondérant dans la régulation de l'expression de la fascine 1, qui est une protéine de liaison à l'actine médiant le phénomène de métastase des tumeurs. La galectine 3 induit également l'expression de la P-sélectine qui participe à la fixation des plaquettes aux leucocytes. L'interaction des leucocytes et des plaquettes mène à l'activation des deux types cellulaires qui engendre alors un déclenchement du processus inflammatoire et l'apparition d'un thrombus. Ainsi, l'implication de la protéine galectine 3 dans différents processus du cancer et de l'inflammation ouvre la voie au développement d'anticorps thérapeutiques pour bloquer les mécanismes par lesquels la protéine galectine 3 favorise ces pathologies.

Ainsi, nous avons démontré ici que la technologie de phage display *in vivo* suivie d'une caractérisation de la cible *in vitro* pouvait être utilisée pour la sélection d'anticorps humains potentiellement utilisables en tant qu'agents thérapeutiques comme par exemple en tant qu'anticorps antagonistes. D'autres anticorps issus de la sélection *in vivo* restent encore à caractériser et permettront dans un futur proche de mettre en valeur d'autres cibles potentielles. Un aperçu des résultats prometteurs obtenus en cytométrie de flux et en immunohistochimie sur les deux modèles précliniques et les endartériectomies humaines après reformatage de ces anticorps est donné en figures annexes (**Figure 54 et Figure 55**).

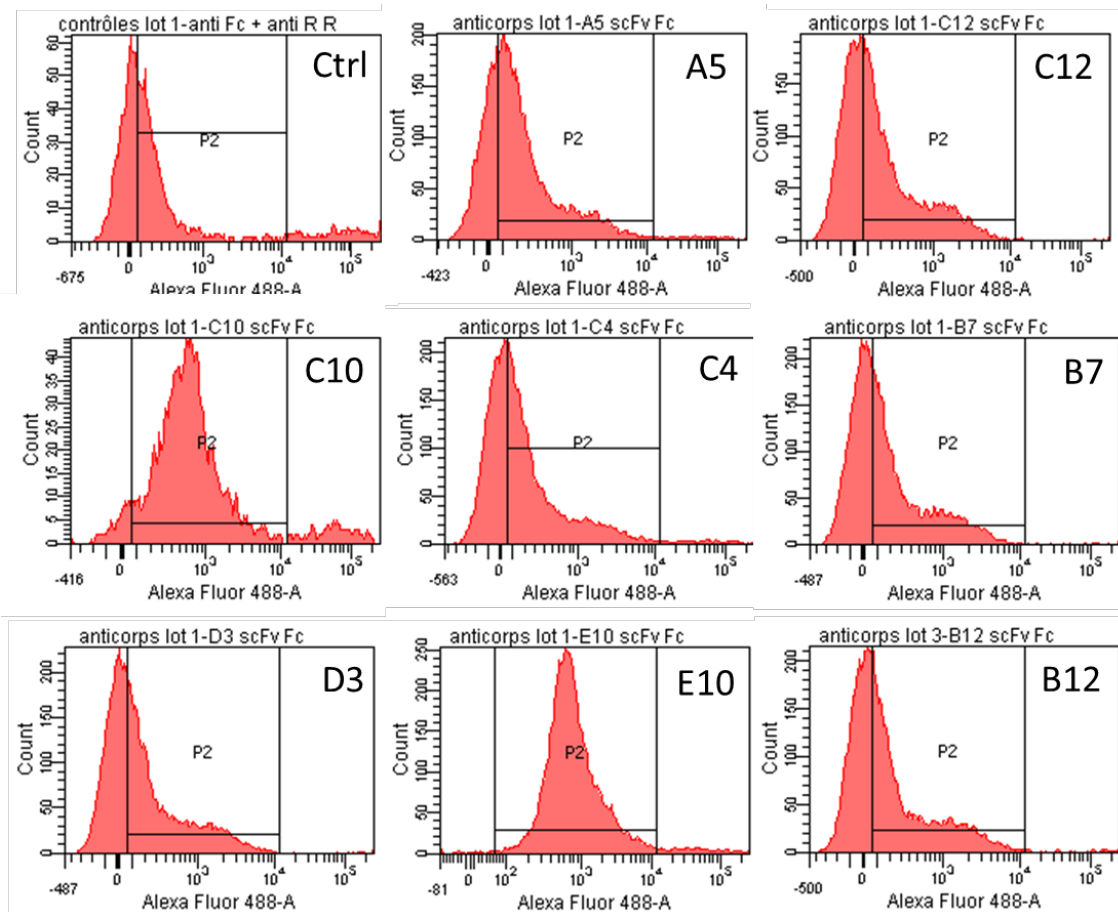


Figure 54 Annexe 1 : test par cytométrie de flux des anticorps reformatés sous format scFv-Fc. Le nombre de billes (count :y-axis) est donné en fonction de l'intensité de fluorescence des scFv-Fc (Alexa488 : x-axis). L'interaction des scFv-Fc avec les protéines athéromateuses a été révélée par l'ajout d'un anticorps lapin anti Fc humain superposé d'un anticorps anti lapin couplé au fluorochrome Alexa488.

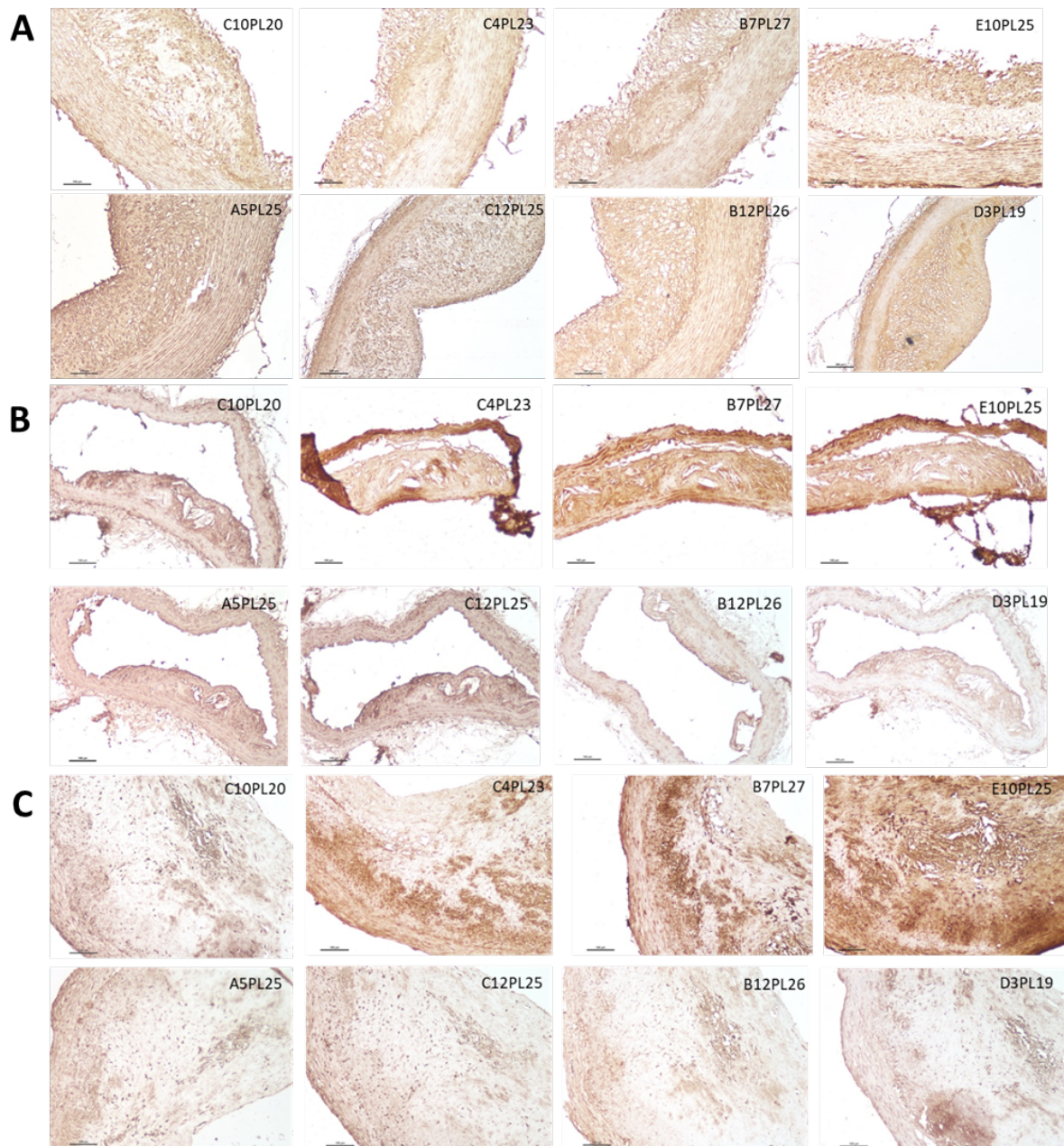


Figure 55 Annexe 2 : immunohistochimie sur coupes de lapin athéromateux (A), de souris ApoE $-/-$ (B) et d'endartériectomies humaines (C). Les anticorps scFv-Fc ont été incubés sur les coupes de tissus athéromateux. L'interaction a été révélée par l'ajout d'un anticorps lapin anti Fc humain couplé à l'enzyme HRP (Horse Radish Peroxidase) suivi de l'ajout du substrat DAB. La fixation est montrée par l'apparition d'un précipité marron. Grossissement x10.

Identification de la galectine 3 en tant
que cible d'un anticorps sélectionné par
phage display *in vivo* et validé *ex vivo*
(Article)

Galectin 3 as a new biomarker of atheroma plaque instability identified by *in vivo* phage display

Audrey HEMADOU¹, Alexandre FONTAYNE², Cyril LORENZATO¹, Jeanny LAROCHE-TRAINEAU¹, Philippe MONDON², Stéphane CLAVEROL³, Stéphane SANCHEZ¹, Florence OTTONES¹, Martine CERUTTI⁴, Gisèle CLOFENT-SANCHEZ¹ and Marie-Josée JACOBIN-VALAT¹

¹ CRMSB, UMR5536 CNRS, INSB, Bordeaux, 33076, France

² LFB Biotechnologies, department of biotherapeutic engineering Lille, 59000, France

³ CGFB, Proteome pole, Bordeaux, 33076, France

⁴ UPS 3044, CNRS, Saint-Christol-Lès-Alès, France

Keywords: Phage display, Antibodies, Flow cytometry, Targets, Imaging

1 Introduction

Atherosclerosis, the main cause of death in western world, is a chronic and inflammatory disease characterized by the build-up of a lipid rich plaque that clogs the arteries. Atherosclerotic lesions result from local accumulation of lipids, immune and non-immune cells and cellular debris¹. Endothelial cells, activated by lipoproteins, express chemokines and adhesion molecules, which help for the recruitment of monocytes. Resident monocytes differentiate into macrophages expressing scavenger receptors and CD36 molecule. These receptors participate to the binding and phagocytosis of oxLDL and lead to the formation of foam cells^{2,3}. Thus, macrophages and foam cells are the most abundant cells in atheroma plaque and the major factor of instability and vulnerability of the plaque.

To improve diagnosis and treatment of the pathology, the identification of new targetable biomarkers is needed. Currently, diagnosis of atheroma plaque is quite invasive and occurs often at advanced stages of the pathogenesis. The *in vivo* discovery of biomarkers will allow a better understanding of the pathogenesis and the development of new tools for imaging modalities and therapeutic purpose. Considering this, we decided to track targets in an hypercholesterolemic rabbit model by *in vivo* phage display using a human single chain fragment variable (scFv) library. After selection and characterization of bound scFv-phages, they were reformatted as scFv-Fc antibodies and co-immunoprecipitated with their targets. We found galectin 3 protein as the target of one scFv-Fc clone (P3 scFv-Fc clone) presenting with the best cross-reactivity against atherosclerotic proteins issuing from two preclinical models and from human carotid biopsies after endarterectomy surgery.

Galectins belong from β -galactoside-specific lectins family. They play a key role in the immune system, inflammation, wound healing and carcinogenesis⁴. Well described in the case of cancer⁵⁻⁸, the role of galectins in cardiovascular diseases has gained a rising interest. Among these galectins, galectin 3 protein is the only chimeric type characterized by one single carbohydrate-recognition domain (CRD) and a flexible amino-terminal tail, composed of a collagen-like structure represented by 7-14 repeats of a 9-amino acid sequence, which are fused onto the CRD to form a pentamer. Galectin 3 protein can be localized at the nucleus as well as the cytoplasm. It can be secreted by an unknown mechanism outside of the cells or bound on cell-surface receptors^{4,9}. In atherosclerosis, galectin 3 protein is overexpressed by macrophages and endothelial cells^{9,10}. It has been shown that galectin 3 protein is implicated in macrocalcification of atheroma plaque¹¹, macrophage activation and monocyte attraction¹². Galectin 3 protein can modulate the oxidative stress by stimulating superoxide production in monocytes and act on the adhesion of monocytes/macrophages to endothelial cells¹³. Galectin 3 protein has been shown to be correlated with plaque instability and vulnerability¹⁰, certainly due to its close interrelation with macrophages and foam cells. An inactivation of galectin 3 gene has shown to remarkably reduce the atheroma thickness in ApoE KO mice^{14,15}. Statin treatment significantly reduced macrophage content of galectin 3, designing this protein as a biomarker of the severity of plaque inflammation¹⁶. In the same study, no such inverse correlation could be established for galectin 1.

Considering the specific binding of P3 scFv-Fc clone to galectin 3 protein as revealed by flow cytometry, ELISA assays and immunohistochemistry on mouse and rabbit models and human tissues, we investigated the

binding of P3 clone *ex vivo* in ApoE KO mice fed with a high cholesterol diet. We observed that P3 scFv-Fc clone was able to bind its target after injection in the atherosclerotic aorta. Thus, P3 scFv-Fc clone could be used as a ligand for functionalization of nanoparticles dedicated to the theranostic of atherosclerosis.

2 Material and methods

2.1 Phage antibody preparation

Individual phage-infected XL1 blue were grown in 96 deep well plates (Greiner Bio one, France) in 500 μ L 2xTY/ampicillin/ 5% glucose supplemented with tetracycline at 10 μ g/mL. After an overnight incubation, 25 μ L of bacterial culture were inoculated with 500 μ L 2xTYGAT and incubated for 3hrs. Phage production was induced by adding 25 μ L of 2TY containing 3.10^8 M13KO7 helper phage particles (Stratagene, France). After infection for 1 h at 37°C, bacteria were pelleted and resuspended in 500 μ L 2x2TYGA supplemented with kanamycin at 40 μ g/mL for last growth at 26°C under rotation (Newbrunswick, Edison, USA). Bacteria were then spun down at 10,000 g, 10 min, and supernatants used immediately for flow cytometry assay.

2.2 Production of recombinant scFv-Fc antibody in HEK-293 FreeStyle and BaculoMab system

2.2.1 Production and purification from HEK-293 cells

ScFv-Fc were produced by transient transfection using the FreeStyle™ 293-F expression system (Invitrogen, France). Expression vectors were prepared by first synthesizing (Invitrogen GeneArt) the cDNA fragment coding the scFv as a linear fragment with optimized codons for Homo sapiens. Each fragment was then subcloned by recombination using Infusion kit in between BamHI and NotI restriction sites of a pCEP4 vector (Invitrogen) containing signal peptide and a human IgG1 Fc fragment. HEK FreeStyle cells were transfected with the purified expression vector according to the supplier instruction with PEI 250 kDa (Sigma-Aldrich) in a 1/2 ratio (DNA/transfection reagent). After 7 days of production at 37°C 8% CO₂, in F17 medium supplemented with 8 mM Gln, the supernatant was harvested, clarified by centrifugation and sterilely filtered at 0.2 μ m prior to dosage by ELISA using FastELISA kit (RD Biotech, France) for human IgG. When scFv-Fc were used as purified molecules, this was performed by a one-step affinity chromatography with HiTrapp FF, protein A (GE Healthcare, France) on AKTA avant 80. The molecules were eluted with 25 mM citrate and dialysed against PBS, sterilely filtered at 0.2 μ m and stored at 4°C prior usage. The quality of the purified scFv-Fc was assessed by SDS-PAGE, size exclusion chromatography on Superdex 200 (GE Healthcare, France) and the endotoxin level was determined by LAL (Lysat d'Amoebocytes de Limule) test.

2.2.2 Production and purification from insect cells

Fusion protein ScFv-Fc was produced using the BaculoMab system. Briefly, the cDNA encoding the P3 scFv was amplified by PCR using the forward and reverse primers: 5'- GCTA CTTAAG GGT GTC CAG TGC CAGGTGCAGCTGCAGCAGTCTGGACCCGG- 3' and 5'- GCTA CGTACG CTTGATTTCCAGCTTGGTGCCGCCT- 3', respectively. The PCR fragment was then inserted into a specific transfer vector in frame with a sequence encoding an IgG1 signal peptide at the 5' end and with a cDNA encoding a human IgG1 Fc domain with 2 extra cysteine residues at the C-terminal end. Sf9 cells were cotransfected by lipofection with the transfer vector and purified viral DNA in the presence of 40 μ L of DOTAP liposomal transfection reagent (Roche)¹⁷. Recombinant viruses were isolated by plaque assay and productive clones were screened by ELISA. The genomic organization of recombinant viruses was controlled by Southern blotting. Sequence of integrated genes was verified after amplification by PCR and sequencing (Eurofins Genomics, Germany). ScFv-Fc was produced by infecting Sf9 cells adapted to grow in serum free medium, with the selected recombinant virus at a multiplicity of infection of 3 plaque forming units (PFU)/cell. Three days post-infection, cell culture supernatant was harvested and scFv-Fc was purified on HiTrap FF, protein A as recommended by the manufacturer (GE Healthcare) on AKTA Purifier. The quality of the purified scFv-Fc was assessed by SDS-PAGE and silver staining, scFv-Fc preparation was sterilized after filtration on 0.22 μ m membrane (Millex GP, Millipore) and stored at 4°C prior usage.

2.3 FACS analysis

2.3.1 Complex plaque formation in the NZW rabbit animal model

All animal experiments were performed in accordance with the Guide for the Care and Use of Laboratory Animals (NIH Publication No. 85-23, revised 1996) and were approved by the ethic committee of Bordeaux (CEEA50). The NZW rabbit model was fed with a high cholesterol diet as previously described¹⁸ and two surgeries (an desendothelialization and an angioplasty) were performed, two months apart.

2.3.2 Rabbit protein extraction

The protein extraction was performed as previously described¹⁹. Briefly, atherosclerotic rabbit proteins were solubilized with a commercially available T-PER lysis buffer (Thermo Fisher Scientific, France) complemented with a protease inhibitors cocktail (Thermo Fischer scientific, France). Solubilization was performed using a Polytron TP-20 Homogenizer 8 (Kinematica, Lucerne, Switzerland). After two centrifugations, at 13,000 g for 45 min at 4°C to discard insoluble material from the supernatant, the protein concentration of every soluble extract was determined using a Bradford assay kit according to the manufacturer's instructions (Thermo Fisher Scientific, France).

2.3.3 Coupling of proteins extracted from atherosclerotic rabbit model to magnetic beads

Fifty micrograms of proteins extracted from rabbit fed with a high fat diet and developing atheroma plaques, as described previously¹⁹, were covalently coupled to 300nm magnetic carboxyl-adenbeads according to the manufacturer's instructions (Ademtech, France). Three batches of coated protein beads per sample were used for reproducibility.

2.3.4 ScFv-phage screening and scFv-Fc validation by flow cytometry

Binding of scFv-phages or scFv-Fc fusion antibodies to atherosclerotic rabbit proteins was determined by flow cytometry. Forty microliters of atherosclerotic rabbit proteins coated on beads were added at 5 µg/mL to 100 µL of scFv-phages or scFv-Fc diluted at 10 µg/mL in PBS buffer and incubated for 3 h at 4°C under rotation. Mouse anti-pVIII (Abcam, France) or rabbit anti human Fcγ specific (Jackson Immunoresearch, USA) primary antibodies were added at 1:1000 or 1:65 dilutions, respectively, and incubated overnight at 4°C under rotation. Alexa 488 labeled anti-mouse antibody (Life technologies, France) and Alexa 488 labeled anti-rabbit antibody (Life technologies, France) were used at 1:40 and 1:30 dilutions, respectively for final scFv-phages or scFv-Fc detection.

2.4 Immunohistochemistry on rabbit, mouse aorta sections and human endarterectomy sections

Human tissue specimens were provided by Pr DUCASSE Eric, vascular surgeon at CHU Pellegrin. Human biopsies were harvested from patients having undergone an endarterectomy surgery after an acute vascular event. All of the clinical interventions took place at CHU Pellegrin Hospital (Bordeaux, France). All work with tissues from human subjects had been approved by the CPP committee (Comité de Protection des Personnes Sud-Ouest et Outre Mer) of Bordeaux and from the Research Ministry in France (Authorization number DC -2016- 2724). The CPP committee waived the need for patient written consent because surgical waste no longer attached to the person is considered "RES NULLIUS". Patients were informed by the clinicians; if they did not express their opposition to research, the de-identified samples were immediately processed and embedded in paraffin. Immunoreactivity of scFv-Fc antibodies was evaluated on paraffin-embedded atheromatous human, mouse and rabbit sections prepared from atheromatous specimens. Each paraffin embedded specimen was deparaffinised and rehydrated. Blocking steps (H₂O₂ blocking and unspecific site blocking with PBS/ 1% BSA/ 0.2% Triton X100) and a retrieval step (TRIS 10mM, EDTA 1mM, 0.05% Tween 20) (pH9) were performed. A supplementary blocking of human sections was performed using 5% goat serum followed by addition of goat

anti human IgG H+L or Fc γ specific diluted at 100 μ g/mL. After washing with PBS/ 1% BSA, slides were then incubated with scFv-Fc antibodies at 10 μ g/mL overnight at 4°C. Specimens were washed with PBS/ 1% BSA/ 0.025% Triton X100. Specimens were then incubated with the secondary HRP - conjugated goat anti human (Fc γ specific) antibody (1:1000) (Jackson ImmunoResearch, USA). Negative control slides were only incubated with the secondary antibody.

2.5 Immunoprecipitation

Atherosclerotic rabbit proteins used for immunoprecipitations were obtained as previously described. Human endarterectomy proteins were extracted and solubilized as atherosclerotic rabbit protein extract protocol. Three micrograms of each scFv-Fc antibody (CN, P3 or E9) or commercial anti Gal3BP antibody (Abcam, France) were captured with 100 μ L of saturated (BSA) Protein G Microbeads (Myltenyi Biotech, France) and incubated for 1 h at 4°C under rotation. Seventy micrograms of atherosclerotic rabbit or human endarterectomy proteins were incubated with 100 μ L of microbeads coupled to antibodies for 2 h at 4°C under rotation. Immunoprecipitations were then passed through a μ column (Myltenyi Biotech, France). After several washings with suitable buffers, elution was performed following two protocols, one by using pre-heated 2 x SDS Laemlli buffer; the other by using acid buffer (Glycine 100mM pH 2,5). Elution with the 2xSDS Laemlli buffer was performed by removing the column from the magnet to retrieve the whole sample (the target bound to the antibody coupled to microbeads). Elution with the acid buffer was performed on the magnet to recover only the target. Then, 2xSDS Laemlli samples were used for subsequent SDS-PAGE analysis or were sent to mass spectrometry facility for further processing. Samples obtained with an acid elution were used for ELISA assays.

2.6 Mass spectrometry analysis

2.6.1 Analysis of CN, P3 and E9 scFv-Fcs immunoprecipitations with atherosclerotic rabbit proteins

2.6.1.1 nLC-MS/MS analysis and database search

CN, P3 and E9 samples in Laemlli buffer were deposited onto SDS (Sodium dodecyl sulfate)-PAGE gel for concentration and cleaning purpose. Separation was stopped once proteins have entered resolving gel. Colloidal blue stained bands were cut out from the gel and treated as described elsewhere²⁰. Briefly, gel pieces were destained and digested with trypsin overnight at 37°C. Peptides were extracted and injected on a Ultimate 3000 nanoLC system (Dionex, Amsterdam, The Netherlands) coupled to a Electrospray Q-Exactive mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Mass spectrometry data were acquired in a data-dependent mode alternating a MS scan survey on m/z 300 – 2000 with HCD (High Collision Dissociation) MS/MS spectra on top 12 ions. Data were searched by SEQUEST through Proteome Discoverer 1.4 (Thermo Fisher Scientific Inc.) against a subset of the 2015.09 version of UniProt database restricted to *mammalian* Reference Proteome Sets (136487 entries). Spectra from peptides higher than 5000 Da (Dalton) or lower than 350 Da were rejected. The search parameters were as follows: mass accuracy of the monoisotopic peptide precursor and peptide fragments was set to 10 ppm and 0.02 Da respectively. Only b- and y-ions were considered for mass calculation. Oxidation of methionines (+16 Da), carbamidomethylation (+57 Da) and propionamidation of cysteine (+46 Da) were considered as variable modifications. Two missed trypsin cleavages were allowed. Peptide validation was performed using Percolator algorithm²¹ and only “high confidence” peptides were retained corresponding to a 1% False Positive Rate at peptide level.

2.6.1.2 Label-Free Quantitative Data Analysis

As shown in supplementary data, Samples 1, 2 and 3 corresponding to CN (S1), P3 (S3) and E9 (S2), respectively, were compared. Raw LC-MS/MS data were imported in Progenesis LC-MS 4.1 (Nonlinear Dynamics Ltd, Newcastle, U.K). Data processing includes the following steps: (i) Features detection, (ii) Features alignment across the 10 sample, (iii) Volume integration for 2-6 charge-state ions, (iv) Normalization on total protein abundance, (v) Import of sequence information, (vi) Calculation of protein abundance (sum of the volume of corresponding peptides). Noticeably, only non-conflicting features and unique peptides were considered for calculation at protein level. Quantitative data were considered for proteins quantified by a minimum of 2 peptides.

2.6.2 P3 scFv-Fc and Gal3BP immunoprecipitations with human endarterectomy proteins

2.6.2.1 nLC-MS/MS analysis and database search

P3 and Gal3BP samples were loaded in Laemmli buffer onto SDS-PAGE 10% acrylamide. After protein migration and colloidal blue staining, bands were cut out from the gel and treated as described elsewhere²⁰. Briefly, gel pieces were destained and digested with trypsin overnight at 37°C. Peptides were extracted and injected on a Ultimate 3000 nanoLC system (Dionex, Amsterdam, The Netherlands) coupled to a LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Mass spectrometry data were acquired in a data-dependent mode alternating a MS scan survey on m/z 300 – 1700 with CID (Collision Induced Dissociation) MS/MS spectra on top 6 ions. Data were searched by SEQUEST through Proteome Discoverer 1.4 (Thermo Fisher Scientific Inc.) against a subset of the 2016.10 version of UniProt database restricted to *mammalian* Reference Proteome Sets (915686 entries). Tolerances were set to 10 ppm and 0.6 Da for precursors and fragments respectively. Only b- and y-ions were considered for mass calculation. Oxidation of methionines and carbamidomethylation of cysteines as fixed modification. Two missed trypsin cleavages were allowed. Peptides were validated at a False Positive Rate of 1% using Percolator²¹.

2.7 ELISA

ELISA plates were coated with proteins issuing from immunoprecipitations or recombinant Galectin3 (Abcam, France), Galectin3BP (Abcam, France), Galectin 1 (Abcam, France), GPαIIbβ3 (Enzyme research laboratory, United Kingdom) or BSA at a concentration of 5 µg/mL in carbonate buffer overnight at 4°C and were blocked with PBS / 5% milk for 1hour on the following day. Commercially available mouse antibodies directed against human Gal3, Gal3BP, Gal1 proteins or P3 scFv-Fc antibody were added to respective wells and incubated at room temperature for 2 h, followed by Horse Radish Peroxidase (HRP)-conjugated detection antibodies. Commercial mouse antibodies and P3 scFv-Fc antibody were diluted at 10 µg/mL and 50 µg/mL, respectively.

2.8 Immunofluorescence staining

For immunofluorescence staining, human biopsies sections were deparaffinised and rehydrated. After a retrieval step (TRIS 10mM, EDTA 1mM, 0.05% Tween 20)(pH9), slides were blocked with (1) PBS/ 1% BSA/ 0.2% Triton X10 followed by (2) 5% goat serum and (3) addition of goat anti human IgG H+L and Fcγ specific diluted at 100µg/mL. P3 scFv-Fc or commercially available mouse anti human Gal3 antibody were added overnight at 4°C. Slides were incubated with P3 scFv-Fc diluted at 50µg/mL in PBS/ 1% BSA and, as controls, secondary antibody alone or commercial mouse anti human Gal3 antibody according to the manufacturer's instructions. Slides were washed six times for 10 min in PBS/ 1% BSA/ 0.025% Triton X100 and the secondary antibody dilutions were applied (Alexa 568 labeled anti-human (1:200) for scFv-Fc or Alexa 488 anti-mouse (1:200) for commercial antibodies). After 1h incubation, slides were washed in PBS 3 times and mounted with Vectashield (VWR, France). Slides were observed with fluorescent microscope (Nikon, France).

2.9 Fluorescence imaging

The chest was opened by thoracotomy, the heart exposed and the right atrium cut. A 30-gauge needle was inserted in the left ventricle. PBS-heparin (50 UI/mL; 2.5 mL (Sanofi Aventis, 238 Vitry-sur-Seine, France)) was inoculated followed by 10 mL PBS. Perfusion was continued with 2 mL of PBS containing a formulation of either P3 scFv-Fc antibody or control human IgG coupled to Alexa647 fluorochrome according to the manufacturer's instructions (Alexa 647 antibody labelling kit, ThermoFisher Scientific, France). After 20 min, mice were again perfused with 5 mL of 4% v/v paraformaldehyde (PFA). Aorta were then removed and embedded in 80 mm petri dish containing 0.8% p/v high-grade, 245 low melting-point agarose. Aorta were imaged with a Fluorescent Macroscope (Leica Microsystem, France), equipped with a dedicated cube for Alexa647 acquisition. The confocal microscopy experiments were performed at the Bordeaux Imaging Center of the Neurosciences Institute of the University of Bordeaux II, France. The technical support of Sébastien Marais and Christel Poujol was appreciated.

3 Results:

3.1 *In vivo* selection of scFv-phages and screening by flow cytometry

ScFv-phages issuing from a fully human library were selected by *in vivo* phage display in an atherosclerotic rabbit. ScFv-phages were isolated from lesional areas located from the aortic arch to the iliac bifurcations in three fractions (endothelial layer, intra-tissular fraction and intra-cellular fraction) during each round of biopanning. To enrich for more specific antibodies, three cycles of selection were performed. Two thousand different scFv-phages recovered after the third round of biopanning were screened against atherosclerotic rabbit proteins for specificity. Overall, around two hundred scFv-phages (10%) reacted with atherosclerotic proteins by binding at least 2-fold over the background noise of WT (Wild-Type) phages in flow cytometry experiments (unpublished data). Two typical examples of experiments highlighting positives scFv-phages (P3, F6, C8 and D4) are shown in **Figure 1A**. Hence, the selection has yielded a panel of antibody candidates specific for atherosclerotic proteins overexpressed in atherogenesis. The E9 clone was chosen as a negative clone for the rest of the experiments.

According to binding validated by immunohistochemistry (IHC) and antibody sequence integrity, fifty scFv among one hundred and forty two were successfully produced in soluble form in HEK293 cells as scFv-Fc fragments. After scFv-Fc engineering, the reactivity was assessed again by flow cytometry against rabbit atherosclerotic proteins (**Figure 1B**). We can observe that P3, F6, C8 and D4 scFv-Fc conserved their reactivity after processing. The “CN” scFv-Fc corresponds to an irrelevant negative control antibody (courtesy of LFB Biotechnologies). For the sake of clarity, only the results of one of the three batches are shown for each flow cytometry experiments, knowing that the results within the three batches are always highly reproducible in each experiment.

3.2 *In vitro* validation of the bioreactivity of reformatted clones by immunohistochemistry assays

ScFv-Fc clones were tested for their capacity to recognize their targets expressed on atheromatous sections. An example of IHC experiments performed with the fifty soluble antibodies is presented in **figure 2**. As shown, P3 scFv-Fc presented the highest specific binding on atherosclerotic sections compared to the other clones (with a strong labelling of the intimal area). C8 scFv-Fc did not show any significant labelling on atherosclerotic sections. F6 and D4 scFv-Fc showed a specific binding on human endarterectomy section in the intimal area, with a higher signal for F6 scFv-Fc. On rabbit aorta sections, a diffused labelling was observed for both scFv-Fc. On mouse aorta sections, D4 scFv-Fc did not show any specificity compared to F6 scFv-Fc. Furthermore, the final aim of the study is the use of *in vivo* selected human antibodies as targeting moieties for diagnosis of atheroma in humans. The best “translatable” antibody is clearly the P3 scFv-Fc with a strong labelling of atheroma in both animal models and human endarterectomy sections. It was thus selected for identification of its target.

3.3 Antigen identification by immunoprecipitation of the target and proteomic analyses

To go further in the characterization of P3 scFv-Fc, we aimed to identify its antigen among the whole panel of proteins extracted from atheroma. In a first approach, atherosclerotic rabbit proteins, previously obtained for flow cytometry screening, were used for immunoprecipitation experiments. Whole Laemmli sample of P3 scFv-Fc clone was compared to CN and E9 negative control scFv-Fc (1) on a SDS-PAGE using a silver staining to visualize the major bands and (2) by a quantitative label-free like analysis. As shown in **Figure S1**, immunoprecipitates of CN and E9 scFv-Fc (IP S1 and S3) showed numerous bands identical to P3 scFv-Fc (IP S2), the strongly stained one at 55kDa being certainly the half scFv-Fv portion (**Figure S1**). The presence of other contaminating bands might be easily explained by the non-specific binding of cytoskeletal or immunoglobulins over-represented in the atherosclerotic extract to antibody-coated magnetic beads within the pool of proteins extracted from atheroma. Thin extra bands could be visualized with P3 scFv-Fc immunoprecipitate at around 70, 48, 45, 20 and 15kDa. For Label-free experiments, raw LC-MS/MS data were compared to mammalian database. As expected, cytoskeletal proteins and immunoglobulins were the most represented hits in the MS/MS spectra data for all scFv-Fc antibodies. However, MS/MS spectra and normalized abundance intensities identified hits in the P3 scFv-Fc list that were clearly distinct from CN and E9 scFv-Fc controls (**Table S1**). Galectin 3 binding protein (Gal3BP) was listed as the first atherosclerotic protein of interest (normalized

abundance at 576 605 839 and a spectral count equal to 94). As reported in the literature, Gal3BP protein is a beta-galactoside binding protein, implicated in cell-cell adhesion or cell-matrix adhesion and often described as being co-immunoprecipitated with its ligand, the Galectin 3 (Gal3) protein. In order to confirm this finding, keeping in mind that rabbit atherosclerotic proteins were essentially compared to human database, we decided to perform immunoprecipitation experiments using human atheromatous proteins extracted from biopsies issuing from endarteriectomies. Targets within atheromatous human proteins were captured with a murine commercial anti-human Gal3BP antibody in parallel to experiments conducted with the P3 human scFv-Fc antibody. The resulting captured immune complexes were subjected to trypsin digestion and subsequent LC-MS/MS analysis. The Gal3BP protein (65kDa) was identified within the immune complexes using either the commercial anti-human Gal3BP antibody (number of peptides equal to 7) (**Table S3**) or the P3 human scFv-Fc antibody (number of peptides equal to 12) (**Table S2**) with, additionally in the P3 sample, the presence of the Gal3 protein (26kDa) (number of peptides equal to 1) (**Table S2**). To complement the analysis and discriminate between Gal3BP and Gal3 proteins, targets eluted from captured immune complexes using commercial murine anti human Gal3BP antibody or P3 scFv-Fc antibody were analysed by ELISA assays. As shown in **figure 3-1**, no difference of OD values was observed adding or not P3 scFv-Fc clone on coated P3 IP. The anti-human Fc secondary antibody itself gives rise to a high OD value. This could be due either to an incomplete elution process between the antibody and its target or most probably due to the heavy presence of human immunoglobulins within the human atheroma extract, hiding the specific binding of P3 antibody on its target. No clear conclusion can be drawn from these first data. But adding the commercial anti-Gal3BP antibody onto coated P3 scFv-Fc IP highlighted a specific binding, suggesting that Gal3BP protein was present in the P3 scFv-Fc immunoprecipitate (**Figure 3-2**). A background noise was noted using the anti-mouse antibody only, most probably linked to inter-species cross-reactivity, considering the large amount of human antibodies contaminating the P3 scFv-Fc immunoprecipitate. These are of course highly present also in the Gal3BP IP, suggesting again a shielding effect, preventing a clear visualization of the binding of P3 scFv-Fc on Gal3BP IP (**Figure 3-3**). Indeed, no difference was observed on coated Gal3BP IP when incubating or not P3 scFv-Fc before adding the anti-human Fc secondary antibody. Another explanation could be that the target of P3 scFv-Fc was not Gal3BP protein as initially inferred from proteomic analyses but rather Gal3 protein immunoprecipitated with its ligand. In that case, the lower representativity of Gal3 protein in the mass spectrometry analysis of P3 scFv-Fc IP (number of peptides = 1) (**Table S2**) could be explained by a lower rate or accessibility of trypsin cleavage sites than in Gal3 BP. These differences in trypsin cleavage susceptibility are indeed one of the limitations of the method, one of their advantages being its sensibility. This is clearly underlined in **Figure 3-4**, which shows here again a poor difference in OD values when adding or not, on the coated Gal3BP IP, the commercial antibody used for immunoprecipitating the target within the protein extract, suggesting that a very low quantity of target is contained in the coated Gal3BP IP but detectable in LC/MS/MS analysis (number of peptides = 7) (**Table S3**).

In a second approach, we thus analyzed the binding of P3 scFv-Fc on recombinant Gal3 and Gal3BP proteins by ELISA assay to discriminate between these two proteins. Commercial antibodies were used in the ELISA assay as positive controls. As shown in **Figure 4**, P3 scFv-Fc bound preferentially Gal3 protein. The specificity of P3 scFv-Fc was further tested against Gal1 protein and other irrelevant proteins such as BSA and α IIb β 3 integrin. P3 scFv-Fc was able to react with recombinant Gal3 protein with minimal cross-reactivity with Gal1 and Gal3BP proteins. One of the explanations of the cross reactivity with Gal 1 protein could be the targeting of the CRD motif shared by all the galectins⁴. One of the ligand of Gal3 protein being integrins such as GP α IIb β 3⁴ and GP α v β 3²², it could explain the low reactivity of P3 scFv-Fc with GP α IIb β 3 integrin. P3 scFv-Fc did not recognize BSA control protein.

To further validate the target on human tissue, we performed co-localisation assays of P3 scFv-Fc with a commercial murine antibody directed against Gal3 protein. Immunofluorescence experiments performed on human endarteriectomy sections (**Figure 5A**) showed that P3 scFv-Fc labelled intimal areas on atheromatous human sections, similar to those recognized by the commercial anti Gal3 antibody. The merge image highlights the co-localization of P3 scFv-Fc and the commercial anti-Gal3 antibody staining. Thus, these *ex vivo* data on atheroma sections confirmed that P3 scFv-Fc recognized an antigenic site of Gal3 protein, which is overexpressed in human atheroma plaque.

3.4 P3 ScFv-Fc label atheroma plaque in ApoE^{-/-} mouse

Next, the binding profile of P3 scFv-Fc was characterized *ex vivo* on aortas of ApoE KO or control (WT) mice. **Figure 5B** describes *ex vivo* fluorescence experiments assessing the binding of P3 scFv-Fc or control human IgG on aortas of ApoE KO or control mice. In comparison with the signal obtained from control human IgG on ApoE KO mouse, fluorescent imaging revealed a significant signal of P3 scFv-Fc. Human IgG control on control mice and P3 scFv-Fc clone on control mouse did not give significant signals when injected.

4 Discussion

Nowadays, there is a rising interest to access and characterize the molecular and cellular components of atheroma plaques that underlie the clinical conditions of stroke and myocardial infarction. It is now well established that the composition more than the narrowing of arteries defines the condition of plaque rupture. From the above, it is clear that specific contrast agents are urgently needed. The specificity can be offered by the use of human antibodies, which have the capacity to recognize a biomarker with a high affinity to improve the targeting for diagnostic purposes of atherosclerosis. .

Phage display has proven to be an efficient tool for the identification of biomarkers for several tissues and cell types. Keller et al. have developed human scFv to target blood and lymphatic cells in cancer ²³ In atherosclerosis, the use of fresh human tissues has allowed the selection of peptides directed against CD100 molecule ²⁴. The first *in vivo* phage display selection was set up in 1996 by Pasqualini and Arap to map endothelial cells in cancer ²⁵. Later on, Arap et al ²⁶, Staquicini et al.²⁷ and Krag et al.²⁸ initiated a human vascular mapping project and found peptides able to recognize tumor cells in patients with cancer. In atherosclerosis, the use of animal models for *in vivo* phage selection has led to the discovery of peptides ²⁹ and antibodies ^{19,30} able to bind a large panel of targets in their microenvironment giving access to atheroma plaque composition. One important limitation of *in vivo* phage display is the identification of the target among overexpressed proteins extracted from atherosclerotic lesions. Unlike peptides that can be compared to proteins database for similarity analysis²⁷, antibodies is a more complex method that requires robust technologies of screening.

After a first selection by a high-throughput flow cytometry assay and further biological validation, fifty scFv among one hundred and forty two were successfully produced in large scale in HEK 293 cells under scFv-Fc format. Their specificity was first confirmed by flow cytometry against atheromatous rabbit proteins. Further immunohistochemistry assays showed a specific labelling on rabbit aorta sections for more than half of them. However, few seem to recognize atheroma sections on both atherosclerotic mouse and rabbit models as well as on human biopsy sections. P3 scFv-Fc showed a specific high labelling of the intimal layer on mouse, rabbit and human atheroma, suggesting that this antibody recognizes a motif conserved between species. We would like to highlight here the importance to obtain translational antibodies able to recognize their target on human tissues in order to be directly transferred from pre-clinical to clinical studies.

To assess the specific binding of P3 scFv-Fc clone in a mouse pre-clinical model, we performed *ex vivo* injections in blood circulation. P3 scFv-Fc clone, previously identified by *in vivo* phage display and *in vitro* screening against atheromatous rabbit proteins, was also found to bind its target expressed on atheroma plaque in ApoE KO mouse, offering a targeted molecular paradigm that warrants the development of novel specific contrast agents functionalized with P3 human antibody for a non-invasive pre-clinical imaging of atherosclerosis.

As underlined above, a drawback of *in vivo* phage display is the myriad of unknown targets, offering concomitantly the advantage of a large panel of targeting possibilities for antibodies. However, following the *in vivo* phage display approach, identification of the target is mandatory both for a better understanding of the pathogenesis and biotherapeutic innovations. In order to pinpoint the target of P3 scFv-Fc within atheromatous proteins extracted from human biopsies, immunoprecipitation and mass spectrometry analyses were conducted. The target of P3 scFv-Fc was identified as a protein of 25-30 kDa, the Galectin 3 protein. This protein, highlighted in recent studies as a new biomarker of atherosclerosis^{12,28,29}, has been initially well described in cancer^{4,30,31}. Galectin 3 has been defined as a signature of the tumor cells. It has been shown to coat the tumoral cell membranes and be involved in metastasis, endothelial recognition and evasion of immune surveillance through the killing of activated T cells in breast cancer ³⁵. In oesophageal and gastric cancers, Galectin 3 protein has been implicated in cell motility with a possible role in the regulation of fascin 1,

an actin binding protein mediating metastasis of the tumor³⁶. In renal cancer, Xu et al. have demonstrated that Galectin 3 is involved in the resistance of tumoral cells against cytotoxic drugs³⁷. Galectin 3 protein is also a strong inducer of P-selectin which interact with the leukocyte PSGL1 (P-selectin glycoprotein ligand 1) to form platelet-leukocyte aggregates³². The interaction between platelets and leukocytes promotes activation of both cell types which is a crucial condition for triggering inflammation, vascular remodelling and thrombosis. The involvement of galectin 3 in different processes in cancer and inflammation has opened the way to develop therapeutics to block the mechanisms by which the protein promotes these pathologies. For example, treatment with GCS-100, a polysaccharide further tested in phase II trial, detached bound Gal-3 from tumor-infiltrating T cells, resulting in an increase in the effector function³⁹. In atherosclerosis, a still unexplored mechanism of galectin 3 protein is its role in macrophage recruitment and their transformation into foam cells, phenomena which are now considered as a signature of plaque instability. Zhu et al. have shown that ¹²⁵I-oxidized LDL and ¹²⁵I-acetylated LDL were actively endocytosed by galectin-3-CHO cells and the incubation with acetyl-LDL led to intracellular accumulation of cholesteryl esters, indicating the role of galectin-3 in endocytosis of AGE-proteins and modified LDLs⁴⁰. Furthermore, knockout mouse for galectin 3 protein have clearly demonstrated a reduction in atheroma progression and plaque formation¹⁴.

All of these studies converge to point out galectin 3 protein as an ideal target for imaging modalities and for developing biotherapeutics to regulate its influence on atherosclerosis by directly targeting the galectin 3 protein or one of its multiple ligands such as galectin 3 binding protein (Gal3BP)⁴¹, fibronectin⁴², integrins GPα1β3⁴ and GPαvβ3²², laminin⁴³ and matrix metalloproteinase as MMP9⁴⁴ which are also molecules overexpressed in atherosclerosis. As P3 scFv-Fc was able to bind its target *ex vivo*, future studies will aim to test the capacity of this antibody, as antagonist, to block one of the mechanisms implicated in the inflammation process of atherosclerosis. Moreover, to increase the affinity of P3, we can consider a maturation step by creating mutations in the CDR3 regions, which are the main actors in the antibody paratope. Another possibility is based on our recent study performed by Third Generation Sequencing on clones issuing from the *in vivo* selection (unpublished data). This study has highlighted sequences related to P3 scFv clone that already presented additional punctual mutations when compared to the germline sequence. The rescue of P3 related scFv could easily be done by custom gene chemical synthesis and may give rise to clones sharing the same clonotype but presenting with a higher affinity for galectin 3 protein because of increased somatic mutations.

We highlighted here that the combination of *in vivo* selection and *in vitro* characterization of the target is a process taking into account the pathological microenvironment for the selection of human antibodies, which is a favourable starting point for developing biotherapeutic molecules more easily transferable to the clinic.

Conflict of interest

The authors declare no conflict of interest.

Financial support

This study was funded by (1) Grant ANR- 13-BSV5-0018 French National Research Agency Program named ATHERANOS, (2) a public grant from the French National Research Agency within the context of the Investments for the Future Program, referenced ANR-10-LABX-57 and named TRAIL.

Author contributions

Audrey HEMADOU conceived and designed the experiments. She wrote the paper.

Alexandre FONTAYNE contributed by his exchange to the content and the writing of the manuscript.

Jeanny LAROCHE-TRAINEAU conceived and designed the experiments, contributed by his exchange to the content and the writing of the manuscript

Philippe MONDON provided the MGHumab phage library and the technical support to use it.

Stéphane CLAVEROL contributed by his exchange to the content and the writing of the manuscript on the proteomic analysis and provided proteomic data.

Stéphane SANCHEZ conceived and designed the experiments.

Florence OTTONES contributed by his exchange to the content and the writing of the manuscript.

Martine CERUTTI contributed by his exchange to the content and the writing of the manuscript.

Gisèle CLOFENT-SANCHEZ contributed by his exchange to the content and the writing of the manuscript and supervised the project.

Marie-Josée JACOBIN-VALAT conceived and designed the experiments and contributed to the writing of the manuscript and supervised the project.

All the authors have read and approved the final manuscript.

References

1. Hansson, G. K. & Libby, P. The immune response in atherosclerosis: a double-edged sword. *Nat. Rev. Immunol.* **6**, 508–519 (2006).
2. Virmani, R., Kolodgie, F. D., Burke, A. P., Farb, A. & Schwartz, S. M. Lessons From Sudden Coronary Death. *Arterioscler. Thromb. Vasc. Biol.* **20**, 1262–1275 (2000).
3. Kunjathoor, V. V. *et al.* Scavenger receptors class A-I/II and CD36 are the principal receptors responsible for the uptake of modified low density lipoprotein leading to lipid loading in macrophages. *J. Biol. Chem.* **277**, 49982–49988 (2002).
4. van der Hoeven, N. W. *et al.* The emerging role of galectins in cardiovascular disease. *Vascul. Pharmacol.* **81**, 31–41 (2016).
5. Ahmed, H. & ALSadek, D. M. M. Galectin-3 as a Potential Target to Prevent Cancer Metastasis. *Clin. Med. Insights Oncol.* **9**, 113–121 (2015).
6. Song, L. *et al.* Galectin-3 in cancer. *Clin. Chim. Acta Int. J. Clin. Chem.* **431**, 185–191 (2014).
7. Eliaz, I. The Role of Galectin-3 as a Marker of Cancer and Inflammation in a Stage IV Ovarian Cancer Patient with Underlying Pro-Inflammatory Comorbidities. *Case Rep. Oncol.* **6**, 343–349 (2013).
8. Ruvolo, P. P. Galectin 3 as a guardian of the tumor microenvironment. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1863**, 427–437 (2016).
9. Novak, R., Dabelic, S. & Dumic, J. Galectin-1 and galectin-3 expression profiles in classically and alternatively activated human macrophages. *Biochim. Biophys. Acta* **1820**, 1383–1390 (2012).

10. Papaspyridonos, M. *et al.* Galectin-3 is an amplifier of inflammation in atherosclerotic plaque progression through macrophage activation and monocyte chemoattraction. *Arterioscler. Thromb. Vasc. Biol.* **28**, 433–440 (2008).
11. Pugliese, G., Iacobini, C., Blasetti Fantauzzi, C. & Menini, S. The dark and bright side of atherosclerotic calcification. *Atherosclerosis* **238**, 220–230 (2015).
12. Falcone, C. *et al.* Galectin-3 plasma levels and coronary artery disease: a new possible biomarker of acute coronary syndrome. *Int. J. Immunopathol. Pharmacol.* **24**, 905–913 (2011).
13. Nachtigal, M., Al-Assaad, Z., Mayer, E. P., Kim, K. & Monsigny, M. Galectin-3 expression in human atherosclerotic lesions. *Am. J. Pathol.* **152**, 1199–1208 (1998).
14. Nachtigal, M., Ghaffar, A. & Mayer, E. P. Galectin-3 Gene Inactivation Reduces Atherosclerotic Lesions and Adventitial Inflammation in ApoE-Deficient Mice. *Am. J. Pathol.* **172**, 247–255 (2008).
15. MacKinnon, A. C. *et al.* Inhibition of galectin-3 reduces atherosclerosis in apolipoprotein E-deficient mice. *Glycobiology* **23**, 654–663 (2013).
16. Sano, H. *et al.* Human galectin-3 is a novel chemoattractant for monocytes and macrophages. *J. Immunol. Baltim. Md 1950* **165**, 2156–2164 (2000).
17. Felgner, P. L. & Ringold, G. M. Cationic liposome-mediated transfection. *Nature* **337**, 387–388 (1989).
18. Deramchia, K. *et al.* By-Passing Large Screening Experiments Using Sequencing as a Tool to Identify scFv Fragments Targeting Atherosclerotic Lesions in a Novel In Vivo Phage Display Selection. *Int. J. Mol. Sci.* **13**, 6902–6923 (2012).
19. Deramchia, K. *et al.* In Vivo Phage Display to Identify New Human Antibody Fragments Homing to Atherosclerotic Endothelial and Subendothelial Tissues. *Am. J. Pathol.* **180**, 2576–2589 (2012).
20. Fessart, D. *et al.* Proteomic remodeling of proteasome in right heart failure. *J. Mol. Cell. Cardiol.* **66**, 41–52 (2014).

21. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S. & MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **4**, 923–925 (2007).
22. Markowska, A. I., Liu, F.-T. & Panjwani, N. Galectin-3 is an important mediator of VEGF- and bFGF-mediated angiogenic response. *J. Exp. Med.* **207**, 1981–1993 (2010).
23. Keller, T. *et al.* Selection of scFv Antibody Fragments Binding to Human Blood versus Lymphatic Endothelial Surface Antigens by Direct Cell Phage Display. *PLoS One* **10**, e0127169 (2015).
24. Luque, M. C. A. *et al.* Phage display identification of CD100 in human atherosclerotic plaque macrophages and foam cells. *PLoS One* **8**, e75772 (2013).
25. Arap, W., Pasqualini, R. & Ruoslahti, E. Cancer treatment by targeted drug delivery to tumor vasculature in a mouse model. *Science* **279**, 377–380 (1998).
26. Arap, W. & Pasqualini, R. The human vascular mapping project. Selection and utilization of molecules for tumor endothelial targeting. *Haemostasis* **31 Suppl 1**, 30–31 (2001).
27. Staquicini, F. I. *et al.* Vascular ligand-receptor mapping by direct combinatorial selection in cancer patients. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 18637–18642 (2011).
28. Krag, D. N. *et al.* Selection of tumor-binding ligands in cancer patients with phage display libraries. *Cancer Res.* **66**, 7724–7733 (2006).
29. Chung, J. *et al.* Discovery of novel peptides targeting pro-atherogenic endothelium in disturbed flow regions -Targeted siRNA delivery to pro-atherogenic endothelium *in vivo*. *Sci. Rep.* **6**, srep25636 (2016).
30. Robert, R. *et al.* Identification of human scFvs targeting atherosclerotic lesions: selection by single round *in vivo* phage display. *J. Biol. Chem.* **281**, 40135–40143 (2006).
31. Hogas, S. *et al.* Potential novel biomarkers of cardiovascular dysfunction and disease: cardiotrophin-1, adipokines and galectin-3. *Arch. Med. Sci. AMS* **13**, 897–913 (2017).
32. Wang, J. *et al.* Novel biomarkers for cardiovascular risk prediction. *J. Geriatr. Cardiol. JGC* **14**, 135–150 (2017).

33. Ebrahim, A. H. *et al.* Galectins in cancer: carcinogenesis, diagnosis and therapy. *Ann. Transl. Med.* **2**, (2014).
34. van den Brûle, F., Califice, S. & Castronovo, V. Expression of galectins in cancer: a critical review. *Glycoconj. J.* **19**, 537–542 (2002).
35. Simone, G. *et al.* Galectin-3 coats the membrane of breast cells and makes a signature of tumours. *Mol. Biosyst.* **10**, 258–265 (2014).
36. Kim, S.-J. *et al.* Galectin-3 increases gastric cancer cell motility by up-regulating fascin-1 expression. *Gastroenterology* **138**, 1035-1045.e1–2 (2010).
37. Xu, Y. *et al.* Galectin-3 inhibition sensitizes human renal cell carcinoma cells to arsenic trioxide treatment. *Cancer Biol. Ther.* **14**, 897–906 (2013).
38. Schattner, M. Platelets and galectins. *Ann. Transl. Med.* **2**, 85 (2014).
39. Demotte, N. *et al.* A galectin-3 ligand corrects the impaired function of human CD4 and CD8 tumor-infiltrating lymphocytes and favors tumor rejection in mice. *Cancer Res.* **70**, 7476–7488 (2010).
40. Zhu, W. *et al.* The role of galectin-3 in endocytosis of advanced glycation end products and modified low density lipoproteins. *Biochem. Biophys. Res. Commun.* **280**, 1183–1188 (2001).
41. Silverman, A. M., Nakata, R., Shimada, H., Sposto, R. & DeClerck, Y. A. A GALECTIN-3-DEPENDENT PATHWAY UPREGULATES INTERLEUKIN-6 IN THE MICROENVIRONMENT OF HUMAN NEUROBLASTOMA. *Cancer Res.* **72**, (2012).
42. Lagana, A. *et al.* Galectin binding to Mgat5-modified N-glycans regulates fibronectin matrix remodeling in tumor cells. *Mol. Cell. Biol.* **26**, 3181–3193 (2006).
43. Van, F. den B., Buicu, C., Sobel, M. E., Liu, F. T. & Castronovo, V. Galectin-3, a laminin binding protein, fails to modulate adhesion of human melanoma cells to laminin. *Neoplasma* **42**, 215–219 (1995).

44. Ortega, N., Behonick, D. J., Colnot, C., Cooper, D. N. W. & Werb, Z. Galectin-3 Is a Downstream Regulator of Matrix Metalloproteinase-9 Function during Endochondral Bone Formation. *Mol. Biol. Cell* **16**, 3028–3039 (2005).

Figure 1. FACS analysis: typical experiments of atherosclerotic protein-binding scFv-phages and soluble scFv-Fc. The number of coated beads (counts:Y-axis) is given as a function of the fluorescence intensity of scFv-phages or scFv-Fc staining (Alexa Fluor 488:X-axis). Protein binding was detected by mouse anti-pVIII antibody plus Alexa Fluor 488-conjugated anti mouse antibody for scFv-phages and rabbit anti human Fc antibody plus Alexa Fluor 488-conjugated anti rabbit antibody for scFv-Fc.

Figure 2. Immunohistochemical assays on atheromatous NZW rabbit, ApoE KO mouse and human endarterectomy (courtesy from Pr DUCASSE) sections of selected scFv-Fc. Selected scFv-Fc were incubated with atherosclerotic rabbit, mouse and human tissue sections. The first line corresponds to the secondary antibody alone. The second line corresponds to the P3 scFv-Fc, the third to the F6 scFv-Fc, the fourth to the D4 scFv-Fc and the fifth to the C8 scFv-Fc. After addition of secondary goat anti-human Fc antibody, coupled to HRP, sections were revealed with DAB substrate kit reagent. The presence of the antigen recognized by the scFv-Fc was indicated by a yellow-brown staining. No staining was observed in the presence of secondary antibody only. Scale bars represent 100µm. Nuclei were counterstained with hematoxylin.

Figure 3. Binding assessments of P3 scFv-Fc and commercial anti human Gal3BP antibody to immunoprecipitations (IP) by ELISA. IP were eluted with a glycine buffer pH 2,2 and coated on ELISA plate. The OD at 405nm is given on the Y-axis. Values represent mean \pm SD (n = 3). Negative controls were anti human Fc antibody and anti mouse antibody alone. ScFv-Fc or commercial Ab binding were revealed via peroxidase-conjugated secondary antibodies. (1) P3 scFv-Fc IP was incubated with or without P3 scFv-Fc and HRP goat anti human Fc antibody. (2) P3 IP was incubated with or without commercial mouse anti human Gal3BP antibody and HRP goat anti mouse antibody. (3) Gal3BP IP was incubated with or without P3 scFv-Fc and HRP goat anti human Fc antibody. (4) Gal3BP IP was incubated with or without commercial mouse anti human Gal3BP antibody and HRP goat anti mouse antibody.

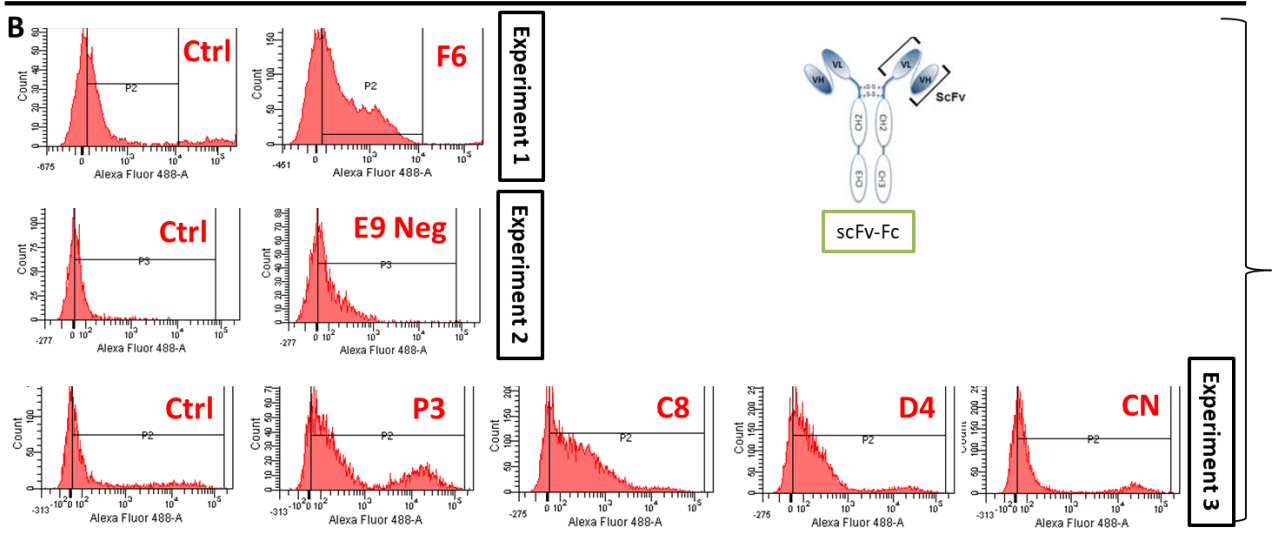
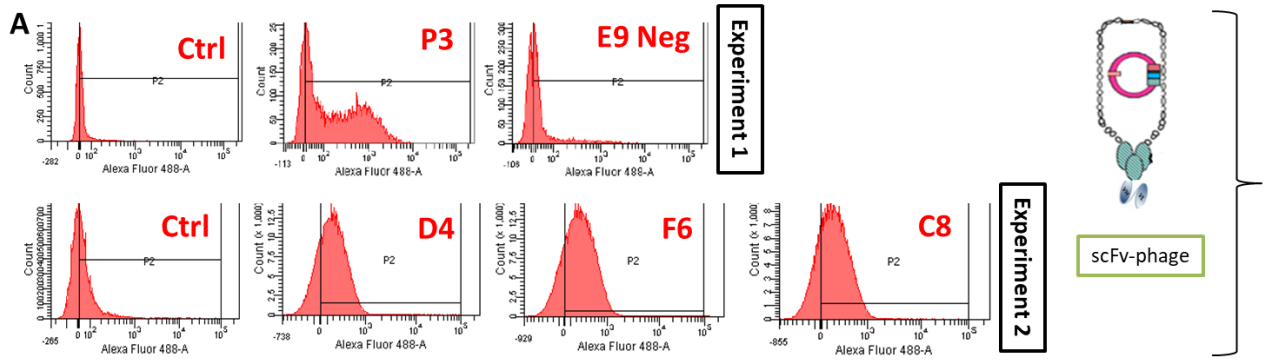
Figure 4. Binding assessments of P3 scFv-Fc and commercial antibodies on recombinant proteins. All recombinant proteins were coated at 5µg/mL. P3 scFv-Fc showed a higher binding on Gal3 than on Gal3BP. Gal1 and GPαIIbβ3 were used to assess the specific binding on Gal3. An irrelevant antigen (BSA) served as control. scFv-Fc binding was detected with a peroxidase-conjugated anti human Fc antibody. Commercial antibodies showing a specific binding on their target were revealed with a peroxidase-conjugated anti mouse antibody. Absorbance was measured at 405nm.

Figure 5. Immunofluorescent staining of human endarterectomy sections and aortas of APOE KO and WT mouse with P3 scFv-Fc or commercial control antibodies. (A) Double immunofluorescence staining of human endarterectomy sections. Alexa-fluor 488 anti mouse antibody was used to reveal the specific binding of the commercial anti human Gal3 antibody. Alexa-fluor 568 anti human

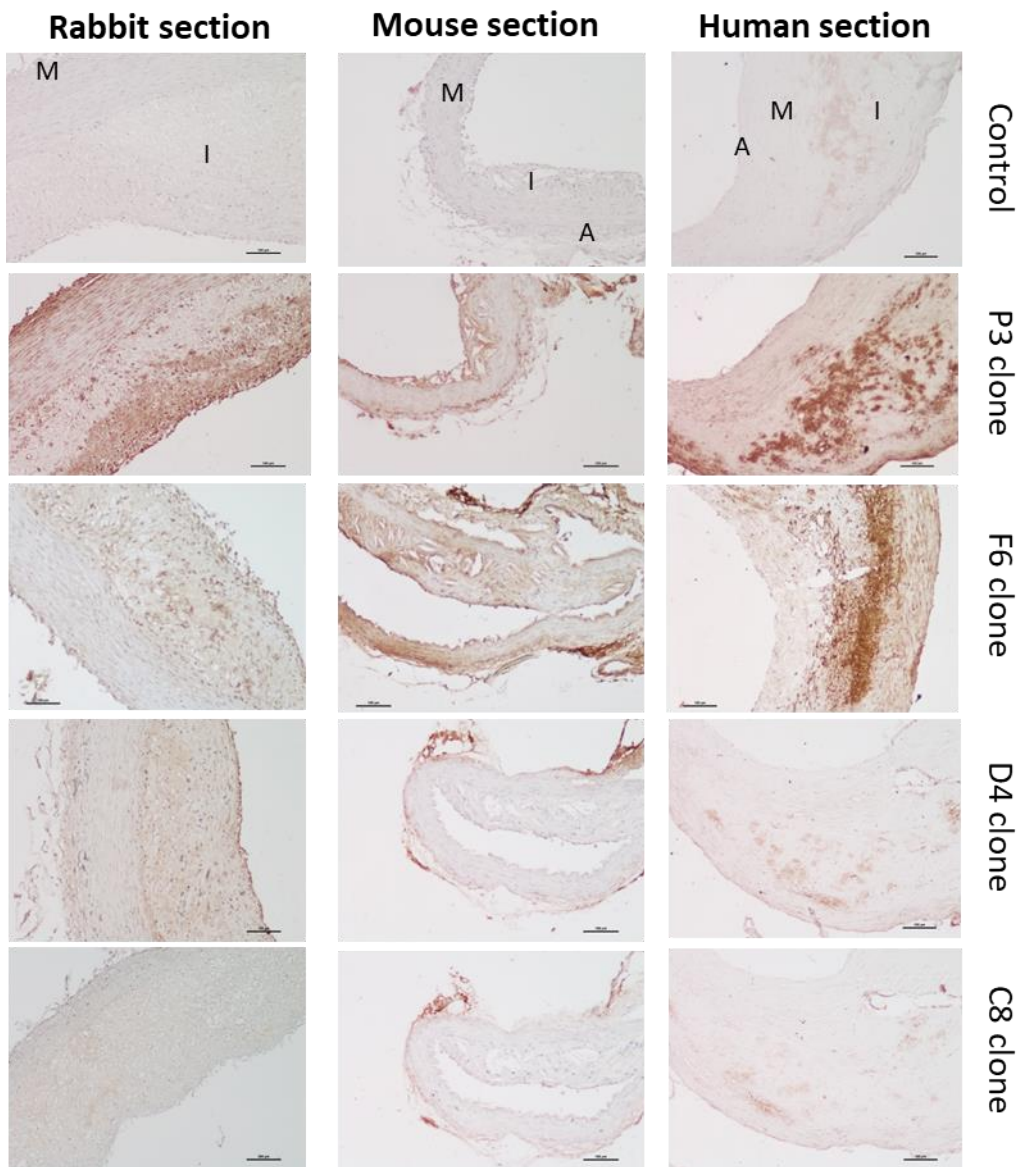
atherosclerosis

antibody was used to reveal the specific binding of P3 scFv-Fc. a) Secondary antibodies alone; (b) Primary antibodies. As seen from co-immunofluorescent staining on human endarterectomy sections, P3 scFv-Fc and anti human Gal3 antibody showed overlapping staining. Size bars: 50 μ m. (B) Fluorescent macroscopy of P3 scFv-Fc after *ex vivo* injection in ApoE KO mouse and WT mouse. A human IgG was used as a negative control antibody. P3 scFv-Fc showed a specific labeling of the atheroma in ApoE KO mouse.

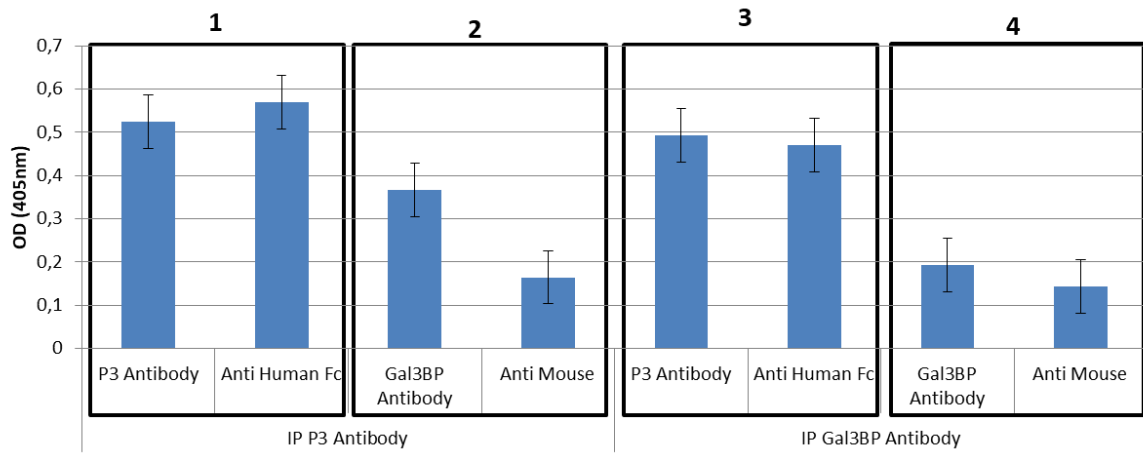
atherosclerosis



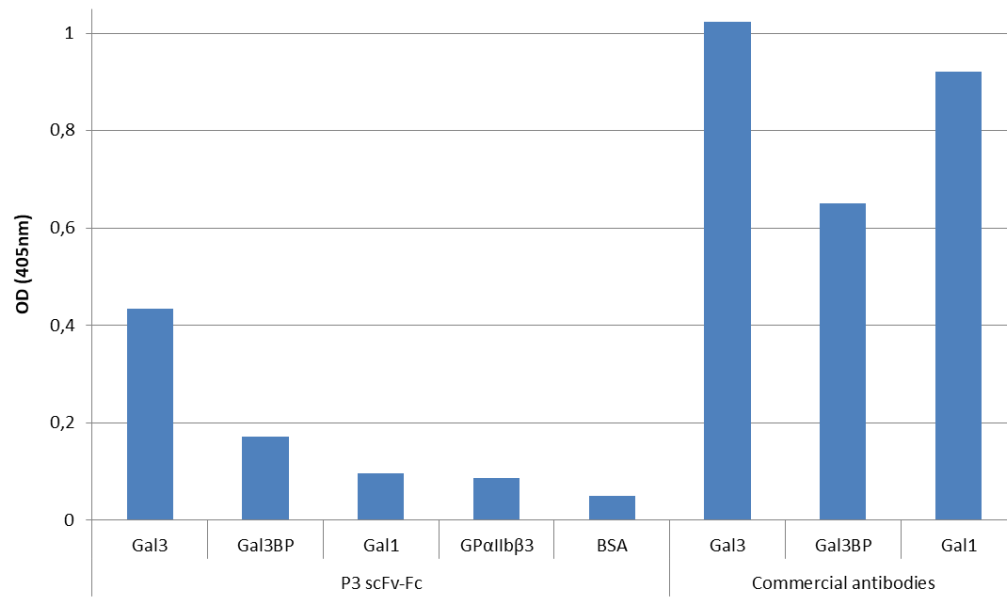
atherosclerosis



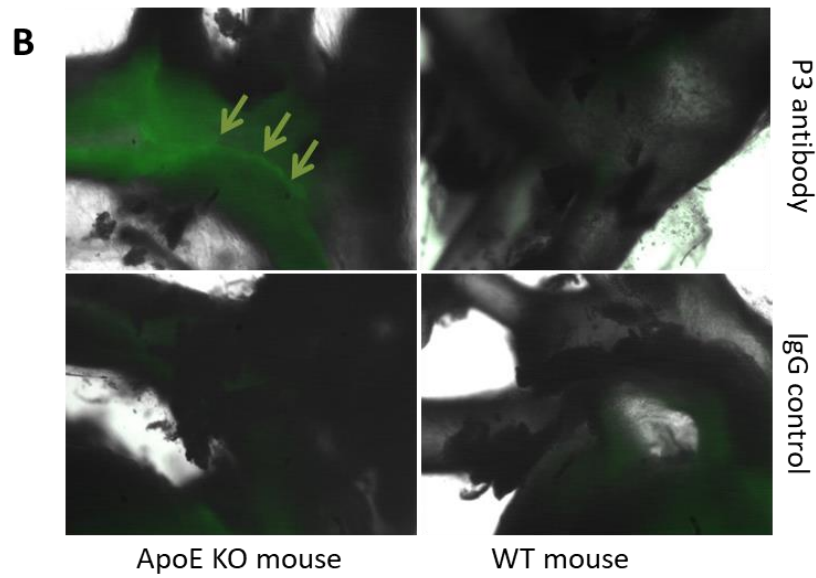
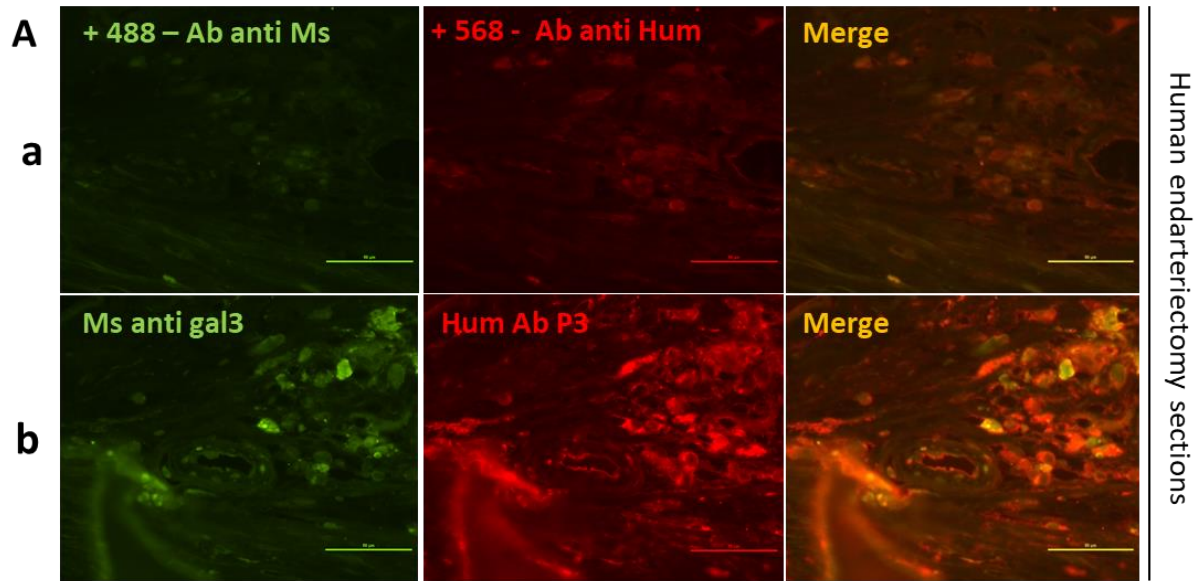
atherosclerosis



atherosclerosis



atherosclerosis



Galectin 3 as a new biomarker of atheroma plaque instability identified by *in vivo* phage display

Audrey HEMADOU¹, Alexandre FONTAYNE², Cyril LORENZATO¹, Jeanny LAROCHE-TRAINEAU¹, Philippe MONDON², Stéphane CLAVEROL³, Stéphane SANCHEZ¹, Florence OTTONES¹, Martine CERUTTI⁴, Gisèle CLOFENT-SANCHEZ¹ and Marie-Josée JACOBIN-VALAT¹

¹ CRMSB, UMR5536 CNRS, INSB, Bordeaux, 33076, France

² LFB Biotechnologies, department of biotherapeutic engineering Lille, 59000, France

³ CGFB, Proteome pole, Bordeaux, 33076, France

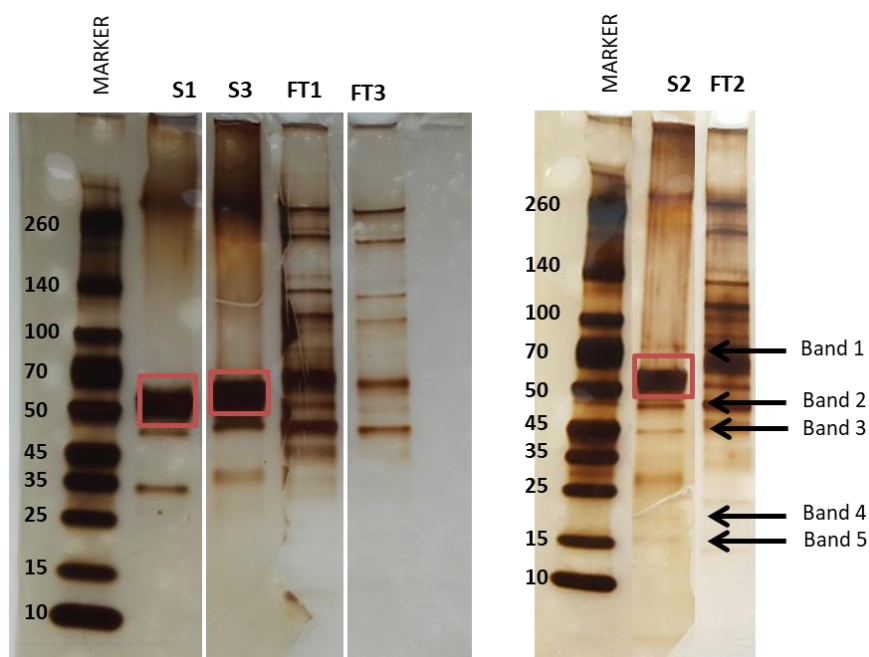
⁴ UPS 3044, CNRS, Saint-Christol-Lès-Alès, France

Keywords: Phage display, Antibodies, Flow cytometry, Targets, Imaging

Supplementary legends

Figure S1. SDS-PAGE analysis of immune complexes. Immunoprecipitations with P3 scFv-Fc (S2) and two negative control clones CN (S1) and E9 (S2) were performed with atherosclerotic rabbit proteins. After elution with 2x SDS laemlli buffer, immune complexes were loaded on polyacrylamide gel and migrated. A silver staining was made. P3 scFv-Fc showed four additional bands at 70, 48, 45, 20 and 15 Kda (Bands 1 – 5) compared to CN and E9 scFv-Fc. The red rectangle shows the band corresponding to the half scFv-Fc portion. The flowthrough of each immunoprecipitation was migrated as controls.

Supplementary figures



Description	Normalized abundance			Spectral counts		
	CN antibody	C3 antibody	E9 antibody	CN antibody	C3 antibody	E9 antibody
MCG124046 OS=Mus musculus GN=Prss1 PE=1 SV=1 - [Q9Z	3 832 023 419	2 266 090 119	2 210 036 561	20	20	13
Ig kappa-b4 chain C region OS=Oryctolagus cuniculus PE=1	1 698 206 618	1 450 755 472	1 830 240 179	52	73	57
Ig gamma chain C region OS=Oryctolagus cuniculus PE=1 S	936 267 806	1 417 748 607	1 312 406 148	48	74	66
Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=	48 809 866	1 388 580 452	47 956 955	1	43	12
Galectin-3-binding protein OS=Homo sapiens GN=LGALS3B	54 751 555	576 605 840	149 769 066	7	94	28
Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=	232 384 089	390 321 828	552 404 901	30	52	59
Heat shock 70 kDa protein 1A OS=Homo sapiens GN=HSPA	12 970 413	362 486 157	73 959 199	2	56	4
Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 P	958 640 549	344 873 441	271 329 505	143	75	52
Actin, aortic smooth muscle OS=Homo sapiens GN=ACTA2	569 139 162	297 414 080	1 220 310 523	68	49	136
Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 f	1 020 102 607	238 210 573	527 198 917	167	102	133
Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT1(1 902 131 677	226 590 112	737 670 404	152	49	110
Uncharacterized protein OS=Oryctolagus cuniculus GN=HSI	821 671 001	209 649 444	303 609 958	4	13	17
Elongation factor 1-alpha 1 OS=Homo sapiens GN=EEF1A1	42 374 648	185 082 205	75 806 506	19	65	36
Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sa	21 971 316	184 685 533	22 047 196	6	27	6
MCG140784 OS=Mus musculus GN=Try10 PE=1 SV=1 - [Q7:	177 322 733	129 174 114	128 473 448	14	7	8
Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens	915 979 933	103 329 417	248 405 153	103	40	61
Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV	84 361 761	101 921 921	118 883 594	11	9	9
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	26 844 934	85 535 691	35 020 225	14	50	23
GTP-binding nuclear protein Ran OS=Homo sapiens GN=RA	5 682 260	72 342 477	11 507 578	4	23	4
Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [E]	10 668 750	71 408 713	9 803 591	-	21	3
Midkine OS=Homo sapiens GN=MDK PE=1 SV=1 - [MK_HU]	7 475 036	67 749 404	6 588 595	-	17	2
Uncharacterized protein OS=Oryctolagus cuniculus GN=VIN	95 854 461	59 300 249	231 215 275	43	38	86
Phosphoglycerate kinase 1 OS=Homo sapiens GN=PGK1 PE=	1 860 250	59 201 935	5 411 683	-	36	4
Fatty acid synthase OS=Homo sapiens GN=FASN PE=1 SV=3	3 951 750	58 354 541	7 126 024	-	57	2
Metalloproteinase inhibitor 1 OS=Homo sapiens GN=TIMP1	8 480 773	56 061 477	2 560 274	-	4	-
40S ribosomal protein SA OS=Homo sapiens GN=RPSA PE=1	10 922 923	54 419 365	11 578 852	3	14	4
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	38 347 874	52 472 575	52 763 654	12	11	13
Guanine nucleotide-binding protein subunit beta-2-like 1 C	23 561 006	51 937 807	27 280 996	16	21	12
Importin subunit beta-1 OS=Homo sapiens GN=KPNB1 PE=	3 907 225	44 845 416	8 282 424	1	26	2
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	31 587 818	42 152 826	46 079 418	8	7	10
Uncharacterized protein OS=Oryctolagus cuniculus GN=GP	128 230 249	42 123 074	127 263 183	26	11	19
T-complex protein 1 subunit beta OS=Homo sapiens GN=C	3 302 095	41 962 880	8 082 055	-	27	3
Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSF	3 176 086	41 929 727	5 382 221	-	28	4

Myosin light polypeptide 6 OS=Homo sapiens GN=MYL6 PE	37 898 426	41 390 405	23 266 244	11	18	9
Claathrin heavy chain OS=Mus musculus GN=Cltc PE=1 SV=1	7 986 883	39 604 434	11 992 964	1	49	11
Biglycan OS=Oryctolagus cuniculus GN=BGN PE=4 SV=1 - [L	70 277 795	38 725 587	83 872 885	19	16	32
Myosin-11 OS=Oryctolagus cuniculus GN=MYH11 PE=2 SV=	11 843 194	37 663 640	26 166 978	6	47	32
Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV	23 505 503	36 914 319	35 090 193	8	11	4
Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV	30 756 643	36 184 776	36 531 571	3	3	3
Uncharacterized protein OS=Oryctolagus cuniculus GN=FLN	32 132 616	36 178 753	69 795 991	10	13	18
40S ribosomal protein S3 OS=Homo sapiens GN=RPS3 PE=1	12 538 898	34 958 497	16 197 284	7	21	19
Uncharacterized protein OS=Oryctolagus cuniculus GN=LO(43 746 003	34 764 281	17 662 211	-	2	-
Peptidyl-prolyl cis-trans isomerase B OS=Homo sapiens GN	6 035 243	34 471 556	11 517 559	1	9	6
Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV	76 339 179	34 453 346	94 287 308	34	30	75
Uncharacterized protein (Fragment) OS=Oryctolagus cunicit	46 252 012	32 080 973	34 287 309	9	7	10
L-lactate dehydrogenase B chain OS=Homo sapiens GN=LD	6 707 234	31 101 184	14 826 541	3	15	12
60S ribosomal protein L10a OS=Homo sapiens GN=RPL10A	2 308 852	29 422 361	6 562 049	-	14	3
Uncharacterized protein OS=Oryctolagus cuniculus GN=PRF	219 100 349	29 326 486	256 082 666	78	25	89
Elongation factor 2 OS=Homo sapiens GN=EEF2 PE=1 SV=4	24 327 225	28 804 792	12 780 680	1	24	6
Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1 - [CLUS_HL	2 178 302	28 498 989	3 473 077	1	11	2
L-lactate dehydrogenase A chain OS=Homo sapiens GN=LD	13 414 563	28 351 211	6 367 218	3	14	5
Profilin-1 OS=Homo sapiens GN=PFN1 PE=1 SV=2 - [PROF1.	9 185 787	28 348 988	10 631 566	2	15	3
Ig kappa chain V region 4135 OS=Oryctolagus cuniculus PE=	33 099 878	28 008 426	35 579 477	8	7	7
Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV	32 936 853	26 656 510	31 599 968	11	9	10
Vitronectin OS=Oryctolagus cuniculus GN=VTN PE=4 SV=1 -	50 026 862	25 590 666	40 550 301	9	14	19
Eukaryotic initiation factor 4A-I OS=Homo sapiens GN=EIF4	3 353 179	24 679 879	6 537 559	1	25	7
Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 S	1 210 239	24 575 757	192 722	-	6	-
T-complex protein 1 subunit alpha OS=Homo sapiens GN=T	1 117 043	23 954 009	1 890 327	-	21	-
Decorin OS=Oryctolagus cuniculus GN=DCN PE=2 SV=1 - [P	24 531 851	23 311 433	32 918 754	4	10	10
Elongation factor 1-gamma OS=Homo sapiens GN=EEF1G P	2 710 586	23 248 906	3 086 641	1	15	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=CSF	7 870 848	22 750 795	9 601 683	4	11	8
Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV=1	8 899 171	22 624 614	18 011 027	6	9	8
T-complex protein 1 subunit delta OS=Homo sapiens GN=C	1 316 977	21 734 864	3 958 443	1	23	5
78 kDa glucose-regulated protein OS=Homo sapiens GN=H:	2 030 045	21 328 498	3 550 575	-	22	2
Lumican OS=Oryctolagus cuniculus GN=LUM PE=4 SV=1 - [C	37 363 536	21 026 229	43 398 316	22	17	28
Uncharacterized protein (Fragment) OS=Oryctolagus cunicit	26 694 138	20 785 731	42 162 518	2	2	3
Fibronectin (Fragment) OS=Oryctolagus cuniculus GN=FN1	21 625 963	20 734 854	29 473 311	12	17	22
Peptidyl-prolyl cis-trans isomerase A OS=Homo sapiens GN	910 612	20 445 975	2 334 763	-	18	1

60S ribosomal protein L12 OS=Homo sapiens GN=RPL12 PE	8 181 293	20 240 145	6 033 618	4	8	5
Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV	19 134 766	18 651 212	17 404 067	12	6	11
SPARC-related modular calcium-binding protein 1 OS=Hom	686 282	18 022 896	2 183 490	-	16	3
Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV	17 172 055	17 228 517	18 900 304	2	3	4
ATP-citrate synthase OS=Homo sapiens GN=ACLY PE=1 SV=	1 030 497	16 756 279	1 428 989	-	22	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=PGI	24 523 955	16 549 680	57 095 742	17	12	42
60S acidic ribosomal protein P0 OS=Homo sapiens GN=RPL	1 266 227	16 185 932	4 029 395	2	12	7
Importin-5 OS=Homo sapiens GN=IPO5 PE=1 SV=4 - [IPO5_	3 717 431	15 909 695	1 961 534	-	19	-
Ubiquitin-like modifier-activating enzyme 1 OS=Homo sapi	1 301 992	15 680 245	867 668	-	10	1
Transcription factor BTF3 homolog 4 OS=Homo sapiens GN	5 313 842	15 531 622	9 075 247	1	3	2
Ig kappa chain V region AH80-5 OS=Oryctolagus cuniculus f	16 650 669	14 465 626	14 334 841	3	4	4
T-complex protein 1 subunit zeta OS=Homo sapiens GN=CC	810 964	14 379 662	829 845	-	13	2
Uncharacterized protein OS=Oryctolagus cuniculus GN=LO(11 018 642	14 187 710	28 615 832	5	6	12
T-complex protein 1 subunit theta OS=Homo sapiens GN=C	1 050 956	14 007 110	1 069 557	-	17	2
Chloride intracellular channel protein 1 OS=Homo sapiens (609 094	13 809 327	1 210 880	2	10	1
Hemoglobin subunit beta-1/2 OS=Oryctolagus cuniculus Gf	9 874 610	13 167 938	13 376 333	3	9	6
T-complex protein 1 subunit eta OS=Homo sapiens GN=CC[2 269 426	13 107 511	2 457 208	-	19	3
Uncharacterized protein OS=Oryctolagus cuniculus GN=C1(7 431 745	13 025 786	9 073 959	2	3	4
Transcription factor BTF3 OS=Oryctolagus cuniculus GN=BT	3 632 604	13 025 521	3 732 835	-	7	3
MCG23377, isoform CRA_b OS=Mus musculus GN=Gm879;	9 959 654	12 780 065	9 238 016	4	11	7
ATP synthase subunit beta, mitochondrial OS=Homo sapier	11 636 706	12 473 675	20 734 885	10	12	17
Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV	15 310 467	12 445 766	13 755 389	8	7	6
D-3-phosphoglycerate dehydrogenase OS=Homo sapiens G	766 173	12 284 548	1 301 128	-	14	-
Protein IGHV4-34 (Fragment) OS=Homo sapiens GN=IGHV4	2 219 925 516	12 258 687	48 985 291	60	4	7
Nascent polypeptide-associated complex subunit alpha OS:	3 205 420	11 621 207	4 761 517	2	3	3
RuvB-like 2 OS=Homo sapiens GN=RUVBL2 PE=1 SV=3 - [RU	974 335	11 377 154	975 773	-	14	-
Ig lambda-3 chain C regions (Fragment) OS=Homo sapiens (7 392 969	10 913 296	11 108 334	7	7	8
Glyceraldehyde-3-phosphate dehydrogenase OS=Oryctolag	6 831 544	10 551 508	22 088 674	3	-	1
Uncharacterized protein OS=Oryctolagus cuniculus GN=PRI	4 698 591	10 337 047	4 822 194	2	16	7
T-complex protein 1 subunit gamma OS=Homo sapiens GN	1 992 804	10 336 380	2 109 606	-	14	-
Heat shock protein HSP 90-alpha OS=Homo sapiens GN=HS	649 122	10 228 555	1 444 691	1	15	1
RuvB-like 1 OS=Homo sapiens GN=RUVBL1 PE=1 SV=1 - [RU	616 416	10 189 213	1 440 276	-	15	1
Serine/threonine-protein phosphatase 2A 65 kDa regulator	1 678 307	10 172 752	2 052 624	-	14	2
Pyruvate kinase PKM OS=Homo sapiens GN=PKM PE=1 SV=	768 046	9 901 116	2 054 864	-	4	-
Exportin-2 OS=Homo sapiens GN=CSE1L PE=1 SV=3 - [XPO2	458 763	9 720 384	741 137	-	16	-

Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [AL	13 133 243	9 651 220	10 138 149	15	16	12
Tubulin beta-4B chain OS=Homo sapiens GN=TUBB4B PE=1	3 802 449	9 614 550	7 051 766	2	4	3
Uncharacterized protein OS=Oryctolagus cuniculus GN=FHI	24 366 181	9 479 983	46 901 886	-	5	-
Microfibril-associated glycoprotein 4 OS=Mus musculus GN	29 222 700	9 370 184	37 361 159	8	3	16
Adenosylhomocysteinase OS=Homo sapiens GN=AHCY PE=	574 439	9 090 889	405 472	1	7	-
Glycogen phosphorylase, liver form OS=Homo sapiens GN=	13 092 740	8 971 662	1 109 972	-	14	1
Serine/arginine repetitive matrix protein 1 (Fragment) OS=	17 258 288	8 871 104	8 404 020	-	9	8
Histone H2B type 1-P OS=Mus musculus GN=Hist1h2bp PE=	2 892 871	7 846 681	3 188 383	5	9	5
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	4 292 307	7 722 456	7 342 763	2	4	2
Cullin-associated NEDD8-dissociated protein 1 OS=Homo si	703 649	7 565 746	449 217	-	11	-
Eukaryotic translation initiation factor 3 subunit F OS=Hom	470 312	7 536 421	1 074 613	-	9	1
14-3-3 protein theta OS=Homo sapiens GN=YWHAQ PE=1 S	4 810 777	7 516 654	1 951 317	1	5	2
Malate dehydrogenase, mitochondrial OS=Homo sapiens G	1 238 111	7 255 723	2 429 251	2	12	3
Creatine kinase B-type OS=Homo sapiens GN=CKB PE=1 SV:	348 106	7 184 918	737 560	-	9	-
C-1-tetrahydrofolate synthase, cytoplasmic OS=Homo sapit	618 977	7 153 209	831 339	-	13	-
Poly(rC)-binding protein 2 OS=Homo sapiens GN=PCBP2 PE	833 809	7 126 107	1 095 932	1	4	-
Ig kappa chain V region 3374 OS=Oryctolagus cuniculus PE=	11 868 407	7 102 239	2 326 935	1	3	2
Nucleoside diphosphate kinase OS=Mus musculus GN=Gm:	1 221 699	6 972 839	1 267 168	1	5	2
T-complex protein 1 subunit epsilon OS=Homo sapiens GN:	410 405	6 951 163	760 756	-	15	-
Voltage-dependent anion-selective channel protein 2 OS=C	2 139 636	6 808 529	3 213 064	-	7	2
Ig kappa chain V-l region Wes OS=Homo sapiens PE=1 SV=1	877 541	6 612 255	1 237 399	1	4	1
WD repeat-containing protein 1 OS=Homo sapiens GN=WD	4 271 133	6 468 818	936 478	-	7	1
Trifunctional purine biosynthetic protein adenosine-3 OS=H	315 140	6 397 665	364 266	-	7	1
Galectin OS=Oryctolagus cuniculus GN=LGALS1 PE=4 SV=1	2 758 231	6 310 278	4 727 570	1	3	3
Heat shock 70 kDa protein 4 OS=Homo sapiens GN=HSPA4	549 216	6 129 618	1 301 696	-	8	1
Elongation factor 1-delta OS=Homo sapiens GN=EEF1D PE=	140 048	6 116 526	110 704	-	4	-
Voltage-dependent anion-selective channel protein 1 OS=C	1 924 630	6 026 035	3 935 729	1	7	4
Heterogeneous nuclear ribonucleoprotein K OS=Homo sapi	592 758	6 021 648	735 444	-	7	1
ATP synthase subunit alpha, mitochondrial OS=Homo sapie	7 828 410	6 018 914	22 588 761	4	8	17
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	3 790 719	5 847 774	4 670 307	2	2	2
Triosephosphate isomerase OS=Homo sapiens GN=TP11 PE:	940 111	5 839 163	609 851	-	7	1
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	6 859 595	5 776 351	14 755 178	8	15	17
Uncharacterized protein OS=Oryctolagus cuniculus GN=LG/	13 248 993	5 692 367	15 023 718	9	9	15
Destrin OS=Homo sapiens GN=DSTN PE=1 SV=3 - [DEST_HL	14 126 507	5 606 600	22 120 023	8	5	15
Histone H2A.J OS=Homo sapiens GN=H2AFJ PE=1 SV=1 - [H	1 856 716	5 421 597	1 582 193	-	2	3

40S ribosomal protein S20 OS=Homo sapiens GN=RPS20 PE	5 164 738	5 053 706	5 212 831	5	7	4
EH domain-containing protein 2 OS=Homo sapiens GN=EHI	4 369 548	4 924 332	10 375 407	2	3	9
Ig heavy chain V-I region V35 OS=Homo sapiens PE=1 SV=1	11 339 142	4 888 187	3 140 039 521	-	-	59
Proliferation-associated protein 2G4 OS=Homo sapiens GN	14 166	4 873 080	95 122	-	5	-
Pre-mRNA-processing factor 19 OS=Homo sapiens GN=PRP	532 221	4 818 620	828 514	-	8	1
60S ribosomal protein L11 OS=Homo sapiens GN=RPL11 PE	1 072 978	4 763 737	2 594 311	1	4	4
Uncharacterized protein OS=Homo sapiens PE=4 SV=1 - [AC	11 244 807	4 722 456	8 329 764	3	3	3
Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 f	19 090 156	4 643 626	8 399 161	6	3	7
Ras-related protein Rab-1A OS=Homo sapiens GN=RAB1A F	8 154 698	4 633 498	10 541 114	5	4	7
Carbonyl reductase [NADPH] 1 OS=Homo sapiens GN=CBR1	232 866	4 632 221	282 462	-	7	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=LO(3 805 688	4 581 348	5 642 379	-	1	2
Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14	18 033 250	4 533 056	8 803 594	8	3	6
Polymerase I and transcript release factor OS=Mus musculi	8 746 683	4 316 883	19 470 055	9	6	12
Heterogeneous nuclear ribonucleoproteins C1/C2 OS=Homo	641 122	4 260 241	2 630 033	-	3	1
Poly [ADP-ribose] polymerase 1 OS=Homo sapiens GN=PAR	610 560	4 157 343	306 164	-	2	-
60 kDa heat shock protein, mitochondrial OS=Homo sapien	970 411	4 062 714	1 295 279	1	6	2
Adenine phosphoribosyltransferase OS=Homo sapiens GN=	216 016	4 049 046	496 150	-	9	-
Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2 - [DS	12 590 329	4 048 264	5 947 941	11	5	3
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	7 773 797	3 930 399	8 100 958	6	4	10
Clusterin OS=Oryctolagus cuniculus GN=CLU PE=2 SV=1 - [C	5 801 117	3 899 119	9 802 286	2	5	10
Mimecan OS=Oryctolagus cuniculus GN=OGN PE=1 SV=1 -	6 694 014	3 745 102	12 109 996	3	3	9
Prolactin-inducible protein OS=Homo sapiens GN=PIP PE=1	5 740 911	3 690 709	2 966 026	2	4	1
Cytosolic acyl coenzyme A thioester hydrolase OS=Homo sa	281 011	3 632 831	418 271	-	8	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=PL1	4 153 205	3 506 834	2 043 398	2	2	1
Microtubule-associated protein 1B OS=Homo sapiens GN=I	876 385	3 470 417	1 951 158	1	1	2
Eukaryotic translation initiation factor 2 subunit 3 OS=Hom	121 866	3 428 097	421 953	-	5	-
Poly(rC)-binding protein 1 OS=Homo sapiens GN=PCBP1 PE	1 819 667	3 404 764	1 683 608	1	5	2
Alanine--tRNA ligase, cytoplasmic OS=Homo sapiens GN=A.	111 689	3 401 363	198 435	-	7	1
Peptidyl-prolyl cis-trans isomerase FKBP4 OS=Homo sapien	66 953	3 399 325	202 679	-	4	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=XPC	625 207	3 383 766	314 944	-	7	-
Ribose-phosphate pyrophosphokinase 1 OS=Homo sapiens	978 697	3 287 417	1 703 948	-	4	1
COP9 signalosome complex subunit 8 OS=Homo sapiens GI	552 722	3 233 750	383 328	-	3	1
Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2 - [H	396 908	3 226 707	867 158	-	3	2
Elongation factor 1-beta OS=Oryctolagus cuniculus GN=EEF	569 551	3 195 279	813 498	-	3	-
Nidogen-1 OS=Homo sapiens GN=NID1 PE=1 SV=3 - [NID1_	10 099 807	3 188 615	324 051 398	-	6	-

Proteasome activator complex subunit 3 OS=Homo sapiens	424 078	3 178 714	410 737	-	4	-
Deoxyribose-phosphate aldolase OS=Homo sapiens GN=DE	159 619	3 129 524	671 000	-	6	-
Proliferating cell nuclear antigen OS=Homo sapiens GN=PC	121 505	3 127 231	134 821	-	6	-
Acetyl-CoA acetyltransferase, cytosolic OS=Homo sapiens C	406 986	3 082 229	3 538 190	-	2	-
Adenylosuccinate lyase OS=Homo sapiens GN=ADSL PE=1 S	117 626	3 026 776	161 466	-	5	-
Ras-related protein R-Ras OS=Mus musculus GN=Rras PE=2	6 767 572	3 018 547	4 026 023	2	1	3
Myosin regulatory light polypeptide 9 OS=Homo sapiens GI	349 927	3 011 460	674 076	-	3	1
Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1 - [CYTA_	7 346 362	2 996 984	615 860	5	4	2
26S proteasome non-ATPase regulatory subunit 13 OS=Hor	80 342	2 968 636	193 267	-	4	-
U1 small nuclear ribonucleoprotein A OS=Homo sapiens GI	1 642 452	2 931 508	4 307 733	2	5	5
Uncharacterized protein (Fragment) OS=Oryctolagus cunicu	3 536 516	2 919 056	4 142 738	3	3	4
Ig mu chain C region secreted form OS=Oryctolagus cunicu	3 365 875	2 900 018	1 322 880	-	4	2
Uncharacterized protein OS=Oryctolagus cuniculus GN=AQ	1 038 958	2 894 011	1 704 460	2	2	2
Transcription intermediary factor 1-beta (Fragment) OS=Hc	154 656	2 800 080	204 848	-	5	-
Synaptic vesicle membrane protein VAT-1 homolog OS=Ho	339 575	2 757 063	658 694	-	3	1
Transgelin-2 (Fragment) OS=Homo sapiens GN=TAGLN2 PE	1 230 417	2 697 648	3 232 168	1	5	3
Vinculin OS=Homo sapiens GN=VCL PE=1 SV=4 - [VINC_HUI	2 103 622	2 652 295	4 725 235	3	8	7
Glutathione S-transferase P OS=Homo sapiens GN=GSTP1 F	11 125	2 616 400	21 543	-	4	-
COP9 signalosome complex subunit 4 OS=Homo sapiens GI	572 309	2 572 155	640 408	-	6	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=PR	4 346 587	2 533 519	8 039 375	5	2	14
Annexin A2 OS=Mus musculus GN=Anxa2 PE=1 SV=2 - [AN	4 961 354	2 516 244	3 743 011	3	5	4
Clathrin light chain A OS=Mus musculus GN=CltA PE=1 SV=:	212 874	2 511 049	1 000 533	-	4	3
Phosphoglycerate mutase 1 OS=Homo sapiens GN=PGAM1	25 830	2 487 150	107 075	-	4	-
Uncharacterized protein (Fragment) OS=Oryctolagus cunicu	3 281 838	2 476 718	10 678 244	5	4	8
Ras GTPase-activating-like protein IQGAP1 OS=Homo sapie	20 542 032	2 468 855	5 516 799	3	4	6
40S ribosomal protein S12 OS=Homo sapiens GN=RPS12 PE	202 638	2 423 916	419 364	-	5	1
Heterogeneous nuclear ribonucleoprotein H OS=Homo sap	243 377	2 328 779	834 963	-	3	-
Transgelin OS=Oryctolagus cuniculus GN=LOC100341632 P	3 424 837	2 284 101	12 741 856	3	2	14
26S proteasome non-ATPase regulatory subunit 11 OS=Hor	173 330	2 230 875	399 941	-	6	-
Hypoxia up-regulated protein 1 OS=Homo sapiens GN=HYC	463 051	2 230 333	54 913	-	5	-
Extracellular superoxide dismutase [Cu-Zn] OS=Oryctolagu	2 143 296	2 183 228	3 794 207	1	3	4
Uncharacterized protein (Fragment) OS=Oryctolagus cunicu	942 719	2 175 463	2 729 467	-	2	2
Uncharacterized protein OS=Oryctolagus cuniculus GN=HSI	3 084 291	2 127 664	8 017 345	2	1	4
40S ribosomal protein S27 OS=Homo sapiens GN=RPS27 PE	2 444 023	2 030 251	2 784 192	3	3	2
Programmed cell death 6-interacting protein OS=Homo saf	121 548	2 022 171	418 467	-	4	1

Uncharacterized protein OS=Oryctolagus cuniculus GN=SEF	1 215 518	2 013 764	1 881 373	-	4	-
Poly(U)-binding-splicing factor PUF60 OS=Homo sapiens GI	34 350	1 988 227	247 163	-	4	-
26S proteasome non-ATPase regulatory subunit 10 OS=Hor	9 265	1 914 597	3 227	-	2	-
Probable carboxypeptidase X1 OS=Homo sapiens GN=CPXI	82 892	1 906 389	28 597	-	4	-
Inorganic pyrophosphatase OS=Homo sapiens GN=PPA1 PE	109 648	1 902 134	69 051	-	2	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=CO	3 073 725	1 887 284	7 137 316	2	7	11
60S ribosomal protein L22 OS=Homo sapiens GN=RPL22 PE	1 356 693	1 852 923	1 772 084	1	3	3
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	1 387 907	1 844 270	2 026 052	2	4	2
Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 - [DCD_H	6 815 921	1 839 887	2 964 378	2	2	2
Puromycin-sensitive aminopeptidase OS=Homo sapiens GN	52 265	1 837 799	149 412	-	6	-
Eukaryotic translation initiation factor 3 subunit A OS=Horr	34 782 695	1 835 112	1 062 842	-	4	-
Ig kappa chain V-I region AU OS=Homo sapiens PE=1 SV=1	2 120 175	1 834 418	698 925	2	-	2
Alcohol dehydrogenase class-3 OS=Homo sapiens GN=ADH	103 330	1 824 942	219 672	-	2	1
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	589 542	1 804 700	2 271 707	-	3	3
Protein S100-A8 OS=Homo sapiens GN=S100A8 PE=1 SV=1	4 062 356	1 787 065	7 498 104	6	2	5
ADP-ribosylation factor 1 OS=Homo sapiens GN=ARF1 PE=1	590 055	1 773 162	1 808 601	-	3	3
Voltage-dependent anion-selective channel protein 3 OS=H	767 064	1 773 044	722 516	-	3	2
Cysteine and glycine-rich protein 2 OS=Homo sapiens GN=C	665 056	1 770 742	928 359	-	4	2
Serine protease HTRA1 OS=Homo sapiens GN=HTRA1 PE=1	900 315	1 763 978	1 891 653	-	5	2
A disintegrin and metalloproteinase with thrombospondin	198 610	1 743 854	14 943	-	3	-
Aldehyde dehydrogenase family 16 member A1 OS=Homo	250 777	1 725 851	236 600	-	3	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=LO	370 478	1 715 988	794 171	1	5	2
Ras-related protein Rab-5B OS=Homo sapiens GN=RAB5B P	2 793 614	1 711 096	4 267 177	2	1	3
Ras-related protein Rab-5C OS=Homo sapiens GN=RAB5C P	1 307 769	1 707 316	3 106 758	2	3	4
CTP synthase 1 OS=Homo sapiens GN=CTPS1 PE=1 SV=2 - [I	248 799	1 701 895	12 481	-	6	-
Proteasomal ubiquitin receptor ADRM1 OS=Homo sapiens	19 993	1 673 120	47 779	-	2	-
Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6	4 432 846	1 672 656	9 246 636	3	2	4
Exportin-7 OS=Homo sapiens GN=XPO7 PE=1 SV=3 - [XPO7	82 171	1 651 302	1 596 299	-	5	-
Apolipoprotein E OS=Oryctolagus cuniculus GN=APOE PE=2	2 566 837	1 639 320	3 513 381	3	3	4
Ras-related protein Rap-1b OS=Homo sapiens GN=RAP1B P	1 658 997	1 634 510	3 515 905	1	2	4
Uncharacterized protein OS=Oryctolagus cuniculus GN=CO	4 265 905	1 616 704	7 510 033	2	2	6
Hemoglobin subunit alpha-1/2 OS=Oryctolagus cuniculus P	1 541 962	1 612 863	3 215 083	1	3	4
Uncharacterized protein OS=Oryctolagus cuniculus GN=ASF	2 228 601	1 612 051	2 044 986	-	1	1
40S ribosomal protein S19 OS=Homo sapiens GN=RPS19 PE	1 406 821	1 609 552	2 573 784	3	2	3
Cytoplasmic dynein 1 heavy chain 1 OS=Homo sapiens GN=	338 371	1 580 046	432 995	-	6	1

26S proteasome non-ATPase regulatory subunit 14 OS=Homo sapiens GN=RPS25 PE=1	273 161	1 557 272	476 578	1	2	1
40S ribosomal protein S25 OS=Homo sapiens GN=RPS25 PE=2	1 236 202	1 522 973	2 808 503	2	2	3
Importin-7 OS=Homo sapiens GN=IPO7 PE=1 SV=1 - [IPO7_	289 690	1 521 652	207 452	-	2	-
Methylosome protein 50 OS=Homo sapiens GN=WDR77 PE=1	93 489	1 520 308	82 683	-	2	-
Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=1	331 266	1 516 792	922 937	-	1	1
Actin-related protein 2 OS=Homo sapiens GN=ACTR2 PE=1	107 877	1 481 094	479 369	-	3	2
Serine-threonine kinase receptor-associated protein OS=Homo sapiens GN=KIF203 PE=1	114 904	1 468 647	56 576	-	4	-
Uncharacterized protein (Fragment) OS=Oryctolagus cuniculus	63 137	1 456 717	283 054	-	3	1
Adenylyl cyclase-associated protein 1 OS=Homo sapiens GN=AC10A1 PE=1 SV=1	719 484	1 450 683	100 252	-	2	1
Dihydropyrimidinase-related protein 3 OS=Homo sapiens GN=HSD17B4 PE=1 SV=1	530 312	1 432 374	1 477 111	-	5	3
Alpha-aminoacidic semialdehyde dehydrogenase OS=Homo sapiens GN=ALDH4L1 PE=1 SV=1	80 539	1 422 910	103 609	-	5	2
Small nuclear ribonucleoprotein-associated protein OS=Oryctolagus cuniculus	478 174	1 414 022	1 193 835	-	2	2
Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=1	1 145 685	1 407 881	5 062 022	3	2	5
Uncharacterized protein (Fragment) OS=Oryctolagus cuniculus	1 851 019	1 407 583	759 296	1	2	2
Alpha-centractin OS=Homo sapiens GN=ACTR1A PE=1 SV=1	98 284	1 384 744	49 509	-	3	-
General vesicular transport factor p115 OS=Homo sapiens GN=VPS35 PE=1 SV=1	13 071	1 337 651	27 446	-	4	-
60S ribosomal protein L23a OS=Homo sapiens GN=RPL23A	234 896	1 334 937	14 452	-	2	-
Exportin-T OS=Homo sapiens GN=XPOT PE=1 SV=2 - [XPOT_	55 160	1 327 126	60 237	-	6	-
Collagen alpha-1(XII) chain OS=Oryctolagus cuniculus GN=C12orf65 PE=1 SV=1	422 227	1 309 720	190 027	-	5	1
Ig kappa chain V-l region AG OS=Homo sapiens PE=1 SV=1	88 009 732	1 298 301	818 803	10	-	1
Uncharacterized protein OS=Oryctolagus cuniculus GN=NTF1	1 174 766	1 268 193	2 249 091	3	3	3
Mitotic checkpoint protein BUB3 (Fragment) OS=Homo sapiens GN=BUB3 PE=1 SV=1	32 082	1 257 652	86 745	-	3	-
Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16	1 061 035	1 251 398	1 227 012	1	3	1
Transitional endoplasmic reticulum ATPase OS=Homo sapiens GN=ATP5F1B PE=1 SV=1	1 855	1 249 798	119 815	-	5	-
Ig kappa chain V region K29-213 OS=Oryctolagus cuniculus	501 642	1 226 263	1 319 732	1	1	1
Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit	576 446	1 218 492	887 652	1	2	2
Tyrosine--tRNA ligase, cytoplasmic OS=Homo sapiens GN=TYRO	10 046	1 181 950	70 921	-	4	-
Flaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 - [FLA2_	2 914 323	1 179 176	1 989 879	3	1	4
Desmocollin-1 OS=Homo sapiens GN=DSC1 PE=1 SV=2 - [DSC1	6 510 986	1 177 487	1 345 542	4	2	-
Cofilin-2 OS=Homo sapiens GN=CFL2 PE=1 SV=1 - [COF2_HI	176 666	1 132 548	224 426	-	2	1
Fructose-bisphosphate aldolase A OS=Homo sapiens GN=ALDO	339 839	1 125 355	776 093	-	1	2
Small nuclear ribonucleoprotein F OS=Homo sapiens GN=SNR	893 396	1 120 323	1 152 549	1	3	1
Uncharacterized protein OS=Oryctolagus cuniculus GN=LOC101928	16 345 067	1 116 352	39 497 511	-	2	-
Calponin-1 OS=Homo sapiens GN=CNN1 PE=1 SV=2 - [CNN1	1 616 778	1 097 606	5 180 347	2	4	6
60S ribosomal protein L23 OS=Homo sapiens GN=RPL23 PE=1	728 947	1 075 029	1 004 262	-	1	2

Uncharacterized protein OS=Oryctolagus cuniculus GN=SFT	38 684	1 071 584	494 805	-	3	1
14-3-3 protein zeta/delta OS=Homo sapiens GN=YWHAZ PE	998 673	1 066 909	1 425 233	1	1	3
Dynein light chain 1, cytoplasmic OS=Homo sapiens GN=DY	202 658	1 048 899	846 507	-	3	3
Serum albumin OS=Oryctolagus cuniculus GN=ALB PE=1 SV	265 081	1 040 096	959 666	-	6	3
Signal recognition particle 9 kDa protein OS=Homo sapiens	481 886	1 010 258	182 565	-	2	1
Uncharacterized protein OS=Oryctolagus cuniculus PE=3 SV	487 871	987 819	1 978 736	-	2	-
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	208 042	984 936	333 120	-	3	1
Eukaryotic translation initiation factor 3 subunit K OS=Hom	252 403	977 708	73 800	-	2	-
Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3 - [DESF	3 722 819	971 554	936 940	5	1	3
Uncharacterized protein OS=Oryctolagus cuniculus GN=LO(20 448	969 129	586 696	1	1	1
Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2 - [HORN_	860 170	963 908	653 995	2	4	1
BRCA1-A complex subunit BRE OS=Homo sapiens GN=BRE I	93 878	961 796	251 568	-	4	-
Ras-related protein Rab-7a OS=Homo sapiens GN=RAB7A P	807 606	947 942	1 381 779	-	-	3
Cytochrome b-c1 complex subunit 1, mitochondrial OS=Ml	419 302	943 414	108 394	-	2	1
Integrin-linked protein kinase OS=Homo sapiens GN=ILK PE	1 103 926	930 273	4 054 745	3	2	8
Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV	326 673	918 930	932 686	-	7	3
Cathepsin D OS=Homo sapiens GN=CTSD PE=1 SV=1 - [CATI	1 008 787	914 033	225 853	3	2	-
Peroxiredoxin-2 OS=Homo sapiens GN=PRDX2 PE=1 SV=5 -	343 785	912 871	194 870	2	4	1
Filamin-A OS=Homo sapiens GN=FLNA PE=1 SV=4 - [FLNA_	89 657	912 584	10 517	-	2	-
Lactylglutathione lyase OS=Homo sapiens GN=GLO1 PE=1	382 020	909 174	1 380 923	-	2	-
Activated RNA polymerase II transcriptional coactivator p1	43 264	909 049	190 757	-	3	-
COP9 signalosome complex subunit 3 OS=Homo sapiens GI	72 644	902 292	198 204	-	2	-
Polyadenylate-binding protein 1 OS=Homo sapiens GN=PAI	273 705	897 886	553 443	-	3	-
S-formylglutathione hydrolase OS=Homo sapiens GN=ESD I	1 173	890 583	81 561	-	2	-
Peptidyl-prolyl cis-trans isomerase-like 1 OS=Homo sapiens	187 932	882 474	66 324	-	2	-
4-trimethylaminobutyraldehyde dehydrogenase OS=Homo	178 837	882 196	81 436	-	3	-
Multifunctional protein ADE2 OS=Homo sapiens GN=PAICS	15 800	881 358	18 289	-	2	-
Omega-amidase NIT2 OS=Homo sapiens GN=NIT2 PE=1 SV:	274 574	870 053	197 888	-	2	-
Stress-70 protein, mitochondrial OS=Homo sapiens GN=HS	199 676	860 675	1 150 822	1	2	3
Aldose reductase OS=Homo sapiens GN=AKR1B1 PE=1 SV=:	224 799	852 995	284 929	-	2	-
Tropomyosin beta chain OS=Oryctolagus cuniculus GN=TP	39 726	841 606	183 228	-	3	-
Proteolipid protein 2 OS=Oryctolagus cuniculus GN=PLP2 P	291 532	831 224	175 582	-	2	-
Transportin-1 OS=Homo sapiens GN=TNPO1 PE=1 SV=2 - [T	78 198	830 722	76 934	-	3	-
Ubiquitin thioesterase OTUB1 OS=Homo sapiens GN=OTUE	1 513 926	827 583	3 367 721	-	2	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=DH	116 166	822 934	173 936	-	3	-

Eukaryotic translation initiation factor 3 subunit M OS=Hor	153 889	822 007	457 065	-	-	3	-
COP9 signalosome complex subunit 6 OS=Homo sapiens GI	12 268	821 925	142 275	-	-	3	-
Ras-related protein Rab-11A OS=Mus musculus GN=Rab11;	323 799	814 138	228 451	1	-	2	-
PDZ and LIM domain protein 7 OS=Homo sapiens GN=PDLI	147 058	809 849	411 387	-	-	2	1
Uncharacterized protein OS=Oryctolagus cuniculus GN=ENI	747 488	801 881	1 880 565	-	-	-	3
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	676 914	795 195	539 882	1	-	3	1
Ras-related C3 botulinum toxin substrate 3 OS=Homo sapie	399 623	794 224	633 098	-	-	2	1
Bifunctional purine biosynthesis protein PURH OS=Homo si	3 792	790 885	13 540	-	-	3	-
Acyl-protein thioesterase 1 OS=Homo sapiens GN=LYPLA1 I	41 726	788 023	105 358	-	-	2	-
ADP/ATP translocase 2 OS=Homo sapiens GN=SLC25A5 PE=	601 535	785 839	1 488 778	-	-	2	3
Importin subunit alpha-1 OS=Homo sapiens GN=KPNA2 PE:	-	775 332	7 640	-	-	3	-
Proteasome subunit alpha type-5 OS=Homo sapiens GN=P5	494 695	764 427	368 357	-	-	1	1
40S ribosomal protein S18 OS=Homo sapiens GN=RPS18 PE	634 894	763 243	1 342 905	1	-	2	3
40S ribosomal protein S5 OS=Homo sapiens GN=RPS5 PE=1	161 430	754 443	258 044	-	-	2	1
Actin-related protein 2/3 complex subunit 4 OS=Homo sap	647 309	742 789	64 028	-	-	2	1
Translation initiation factor eIF-2B subunit alpha OS=Homo	47 585	742 474	50 705	-	-	2	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=LO(456 989	726 852	646 484	-	-	2	2
Keratin, type II cytoskeletal 4 OS=Homo sapiens GN=KRT4 f	3 422 045	714 515	37 757 197	-	-	-	14
U2 small nuclear ribonucleoprotein A' (Fragment) OS=Hom	445 776	683 658	25 306 476	-	-	2	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=MA	82 810	680 926	322 709	-	-	2	1
Neutral alpha-glucosidase AB OS=Homo sapiens GN=GANA	3 740 580	671 278	3 735 164	-	-	2	-
26S protease regulatory subunit 10B OS=Homo sapiens GN	188 505	667 456	248 040	-	-	2	1
Junction plakoglobin OS=Homo sapiens GN=JUP PE=1 SV=3	2 606 092	663 229	318 512	3	-	3	1
Eukaryotic translation initiation factor 3 subunit E OS=Hom	63 773	656 549	41 125	-	-	3	-
CD97 antigen OS=Homo sapiens GN=CD97 PE=1 SV=4 - [CD	5 649	638 686	12 831	-	-	3	-
Phenylalanine--tRNA ligase alpha subunit OS=Homo sapien	314 995	637 940	157 661	-	-	2	-
NHP2-like protein 1 OS=Homo sapiens GN=NHP2L1 PE=1 S)	327 428	631 653	146 233	-	-	3	1
Vacuolar protein sorting-associated protein 35 OS=Homo s	108 403	621 588	28 205	-	-	3	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=DP'	202 293	618 868	479 344	-	-	1	2
Heterogeneous nuclear ribonucleoprotein M OS=Homo sar	48 542	613 235	15 109	-	-	3	-
Ig kappa chain V-I region EU OS=Homo sapiens PE=1 SV=1 -	239 280	612 105	191 870	-	-	2	1
Profilin-1 OS=Mus musculus GN=Pfn1 PE=1 SV=2 - [PROF1_	629 424	607 066	981 010	1	-	2	2
Uncharacterized protein OS=Oryctolagus cuniculus GN=HIV	53 180	606 607	145 723	-	-	2	-
Eukaryotic translation initiation factor 5B OS=Homo sapien	52 400	603 902	78 653	-	-	4	-
Hemoglobin subunit beta OS=Homo sapiens GN=HBB PE=1	873 199	592 877	435 640	-	-	1	1

Uncharacterized protein OS=Oryctolagus cuniculus GN=MF	294 584	592 360	494 390	1	1	-
Tissue factor pathway inhibitor OS=Homo sapiens GN=TFPI	291 186	588 656	476 157	-	2	1
Glycylpeptide N-tetradecanoyltransferase 1 OS=Homo sapi	149 018	586 132	185 033	-	2	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=UB	10 955	581 073	33 005	-	3	-
Spondin-1 OS=Homo sapiens GN=SPON1 PE=1 SV=2 - [SPOI	145 845	576 637	409 966	-	2	2
Myosin regulatory light chain 12A OS=Homo sapiens GN=N	117 875	575 202	251 492	-	2	2
Small glutamine-rich tetraatricopeptide repeat-containing p	65 097	571 485	6 588	-	2	-
Tubulin beta-2B chain OS=Homo sapiens GN=TUBB2B PE=1	14 579	566 235	358 003	-	2	2
Replication protein A 32 kDa subunit OS=Homo sapiens GN	88 411	558 067	74 240	-	2	-
Ran GTPase-activating protein 1 OS=Homo sapiens GN=RAI	6 880	540 985	47 944	-	2	-
Eukaryotic translation initiation factor 3 subunit L OS=Hom	53 065	539 221	93 008	-	3	-
Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13	983 062	531 696	48 801 236	-	-	22
Eukaryotic translation initiation factor 3 subunit B OS=Mus	52 514	510 095	80 116	-	2	-
Complement component C8 gamma chain OS=Oryctolagus	550 039	507 076	915 979	1	-	2
Transaldolase OS=Homo sapiens GN=TALDO1 PE=1 SV=2 -]	95 814	505 984	16 472	-	2	-
Adenylosuccinate synthetase isozyme 2 OS=Homo sapiens	23 664	493 880	42 352	-	2	-
Far upstream element-binding protein 2 OS=Homo sapiens	56 049	491 441	107 781	-	2	-
Coatmer subunit delta OS=Homo sapiens GN=ARCN1 PE=:	74 660	487 490	37 812	-	2	-
Actin-related protein 2/3 complex subunit 1B OS=Homo sa	1 118	472 691	3 595	-	3	-
14-3-3 protein gamma OS=Homo sapiens GN=YWHAG PE=]	683 393	467 211	848 944	-	1	1
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	440 968	464 458	771 409	-	3	2
Collagen alpha-2(IV) chain OS=Homo sapiens GN=COL4A2 F	7 333	459 037	25 493	-	2	-
Cystatin-B OS=Homo sapiens GN=CSTB PE=1 SV=2 - [CYTB_	355 683	459 014	853 351	-	1	1
COP9 signalosome complex subunit 1 OS=Mus musculus GI	25 595	458 446	8 822	-	2	-
AP-1 complex subunit beta-1 OS=Homo sapiens GN=AP1B1	10 496	444 961	253 089	-	3	-
Microfibrillar-associated protein 2 OS=Homo sapiens GN=IA	-	444 148	-	-	2	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=MV	392 439	441 601	2 934 295	-	-	7
Valine--tRNA ligase OS=Homo sapiens GN=VARS PE=1 SV=4	7 737	440 957	-	-	2	-
26S protease regulatory subunit 7 OS=Homo sapiens GN=P	62 218	435 266	296 213	-	3	2
COP9 signalosome complex subunit 5 OS=Homo sapiens GI	269 136	427 688	258 696	-	2	-
Probable ATP-dependent RNA helicase DDX6 OS=Homo sa[123	421 894	1 036	-	2	-
Ras-related protein Rab-2A OS=Homo sapiens GN=RAB2A F	405 000	421 370	1 061 003	-	2	2
Transketolase OS=Homo sapiens GN=TKT PE=1 SV=3 - [TKT.	1 444	420 630	9 645	-	3	-
Ribonuclease inhibitor OS=Homo sapiens GN=RNH1 PE=1 S	11 379	417 804	23 229	-	3	-
26S protease regulatory subunit 6A OS=Homo sapiens GN=	80 285	416 383	66 508	-	2	-

Glutamine--fructose-6-phosphate aminotransferase [isome	440 625	415 105	83 379	-	3	-
Protein-glutamine gamma-glutamyltransferase E OS=Homc	661 224	411 597	185 396	1	-	1
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	369 309	404 547	1 551 820	-	2	1
Eukaryotic translation initiation factor 3 subunit C OS=Hom	3 597	398 866	-	-	2	-
Serine protease 23 OS=Homo sapiens GN=PRSS23 PE=1 SV=	8 159	398 515	230 617	-	2	-
Leukotriene A-4 hydrolase OS=Homo sapiens GN=LTA4H Pf	38 483	392 577	21 183	-	3	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=GN	271 797	390 812	1 551 846	2	1	4
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	593 569	389 167	1 929 934	2	-	3
Prohibitin (Fragment) OS=Homo sapiens GN=PHB PE=1 SV=	50 777	376 723	850	-	2	-
Small proline-rich protein 3 OS=Homo sapiens GN=SPRR3 P	60 445	372 386	3 019 823	-	-	2
ATP-dependent 6-phosphofructokinase OS=Oryctolagus cu	976 715	371 549	4 062 233	1	1	6
Zymogen granule protein 16 homolog B OS=Homo sapiens	313 215	370 845	10 351	1	1	-
Phosphoglucomutase-1 OS=Homo sapiens GN=PGM1 PE=1	27 112	356 161	21 257	-	2	-
40S ribosomal protein S14 OS=Homo sapiens GN=RPS14 PE	97 605	338 316	246 059	-	2	1
Keratin, type II cytoskeletal 78 OS=Homo sapiens GN=KRT7	479 532	321 520	1 058 082	1	-	3
Secreted frizzled-related protein 1 OS=Homo sapiens GN=S	-	318 724	1 566	-	2	-
Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE	2 210 226	317 996	113 629	3	-	-
Vacuolar protein sorting-associated protein VTA1 homolog	1 820	316 645	3 927	-	2	-
Apolipoprotein A-I OS=Oryctolagus cuniculus GN=APOA1 P	376 272	306 904	379 675	-	3	1
Lysine--tRNA ligase OS=Homo sapiens GN=KARS PE=1 SV=3	71 272	282 166	75 814	-	3	-
26S protease regulatory subunit 6B OS=Homo sapiens GN=	1 291	281 636	13 628	-	2	-
Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 OS=	93 344	281 238	197 218	-	2	1
Uncharacterized protein OS=Oryctolagus cuniculus GN=NO	694 331	274 841	2 157 518	-	-	4
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	133 387	272 293	224 333	-	2	1
Exportin-5 OS=Homo sapiens GN=XPO5 PE=1 SV=1 - [XPO5	443 301	259 288	84 623	-	2	-
Peroxidase homolog OS=Homo sapiens GN=PXDND PE=1 SV	17 530	258 197	99 642	-	4	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=CD,	414 048	249 195	360 088	-	2	1
Myosin light chain kinase, smooth muscle OS=Oryctolagus	138 498	242 825	693 723	-	2	3
Uncharacterized protein OS=Oryctolagus cuniculus GN=ICA	38 957 578	241 338	2 604 720	2	-	-
26S proteasome non-ATPase regulatory subunit 6 OS=Hom	-	237 311	18 878	-	2	-
Eukaryotic translation initiation factor 2A OS=Homo sapien	25 346	233 422	11 474	-	2	-
Tetranectin OS=Homo sapiens GN=CLEC3B PE=1 SV=3 - [TE	697 684	233 116	548 509	2	-	2
Aminoacyl tRNA synthase complex-interacting multifuncti	66 418	232 072	6 424	-	2	-
Ran-specific GTPase-activating protein OS=Homo sapiens C	75 713	231 848	161 113	-	2	-
Matrix Gla protein OS=Mus musculus GN=Mgp PE=3 SV=1 -	43 368	228 650	188 505	-	2	1

Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2 - [SP	440 067	224 962	287 671	-	2	1
Band 4.1-like protein 2 OS=Homo sapiens GN=EPB41L2 PE=	70 831	220 986	267 885	-	3	-
Protein Ighv1-67 OS=Mus musculus GN=Ighv1-67 PE=4 SV=	852 969	214 473	1 911 146 364	-	-	13
Gamma-glutamylcyclotransferase OS=Homo sapiens GN=G	300 412	208 039	276 469	-	2	-
Integrin beta-1 OS=Homo sapiens GN=ITGB1 PE=1 SV=2 - [I	104 858	203 497	127 334	-	2	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=FAI	154 667	197 583	1 162 118	1	1	4
Caspase-14 OS=Homo sapiens GN=CASP14 PE=1 SV=2 - [CA	4 324 883	197 299	530 570	3	-	1
Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=	190 933	193 439	551 772	-	2	2
Annexin A1 OS=Homo sapiens GN=ANXA1 PE=1 SV=2 - [AN	68 113	192 530	3 914 875	-	-	5
Arginase-1 OS=Homo sapiens GN=ARG1 PE=1 SV=2 - [ARGI	719 404	184 082	4 840	2	-	-
Ras-related protein Rab-14 OS=Homo sapiens GN=RAB14 P	14 836	183 621	246 162	-	2	1
Histone-binding protein RBBP7 OS=Homo sapiens GN=RBB	99 708	179 362	8 028	-	2	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=LPF	18 214	170 244	584 337	-	1	2
40S ribosomal protein S17-like OS=Homo sapiens GN=RPS1	82 761	169 235	65 892	-	2	1
Uncharacterized protein OS=Oryctolagus cuniculus GN=SU	722 865	168 488	1 006 601	2	-	1
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	184 215	167 882	499 134	1	-	1
Large proline-rich protein BAG6 OS=Homo sapiens GN=BAC	29 033	162 177	26 382	-	2	-
Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1	983 344	157 138	1 711 775	1	-	1
Tetrapeptide repeat protein 38 OS=Homo sapiens GN=	58 572	150 197	103 525	-	2	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=SFX	78 731	148 282	973 645	-	-	3
Uncharacterized protein OS=Oryctolagus cuniculus GN=RPI	11 008	142 661	216 860	-	1	1
Protein RCC2 OS=Homo sapiens GN=RCC2 PE=1 SV=2 - [RC	-	142 583	-	-	2	-
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	80 628	131 640	1 562 808	-	2	5
Uncharacterized protein OS=Oryctolagus cuniculus GN=LM	1 960	130 387	65 981	-	1	1
Isocitrate dehydrogenase [NADP], mitochondrial OS=Homo	74 978	124 938	328 228	-	1	3
Threonine--tRNA ligase, cytoplasmic OS=Homo sapiens GN	-	119 150	27	-	2	-
Uncharacterized protein OS=Oryctolagus cuniculus GN=ETF	205 759	117 249	493 228	-	-	2
40S ribosomal protein S8 OS=Homo sapiens GN=RPS8 PE=1	5 931	116 973	192 226	-	1	1
GDP-L-fucose synthase OS=Homo sapiens GN=FTSA3 PE=1	-	110 988	-	-	2	-
Dynammin-1-like protein OS=Homo sapiens GN=DNM1L PE=	1 979	108 679	45 502	-	1	1
Prelamin-A/C OS=Homo sapiens GN=LMNA PE=1 SV=1 - [LA	63 351	108 049	143 560	1	2	1
Endoplasmic (Fragment) OS=Homo sapiens GN=HSP90B1 P	82 601	98 177	64 276	-	1	1
Calponin-3 OS=Homo sapiens GN=CNN3 PE=1 SV=1 - [CNN	45 414	79 210	570 309	-	-	2
Uncharacterized protein OS=Oryctolagus cuniculus GN=ESY	20 183	71 108	49 631	-	1	1
Alpha-parvin OS=Homo sapiens GN=PARVA PE=1 SV=1 - [J3	94 967	60 841	1 441 006	1	-	5

Filamin-B OS=Homo sapiens GN=FLNB PE=1 SV=2 - [FLNB_f	66 019	60 564	28 686	-	2	-
Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV	191 303	47 431	1 663 127	-	-	4
60S ribosomal protein L24 OS=Homo sapiens GN=RPL24 PE	7 482	41 739	568 874	-	1	2
Ras suppressor protein 1 OS=Homo sapiens GN=RSU1 PE=1	48 448	38 314	572 280	-	-	2
Uncharacterized protein (Fragment) OS=Oryctolagus cunic	46 443	26 689	447 920	-	-	2
Annexin A1 OS=Oryctolagus cuniculus GN=ANXA1 PE=1 SV-	15 804	25 479	167 777	-	1	1
Uncharacterized protein OS=Oryctolagus cuniculus GN=EM	131 186	17 343	130 763	1	-	2
Small proline-rich protein 2E OS=Homo sapiens GN=SPRR2I	40 718	15 866	1 255 419	-	-	3
Uncharacterized protein OS=Oryctolagus cuniculus GN=TPF	31 291	11 372	441 483	-	-	2
EH domain-containing protein 3 OS=Mus musculus GN=Ehc	10 081	8 070	585 006	-	-	2

Accession	Description	# Unique Peptides	# Peptides	# PSMs
P01857	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1 - [IGHG1_HUMAN]	15	28	218
A0A075B6N8	Ig gamma-3 chain C region (Fragment) OS=Homo sapiens GN=IGHG3 PE=1 SV=1 - [A0A075B6N8_HUMAN]	6	22	147
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6 - [K2C1_HUMAN]	37	43	136
P01859	Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1 SV=2 - [IGHG2_HUMAN]	8	20	141
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6 - [K1C10_HUMAN]	29	38	119
P60709	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 - [ACTB_HUMAN]	10	27	123
P01834	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1 - [IGKC_HUMAN]	9	9	92
P68032	Actin, alpha cardiac muscle 1 OS=Homo sapiens GN=ACTC1 PE=1 SV=1 - [ACTC_HUMAN]	10	26	116
A0A075B6K9	Ig lambda-2 chain C regions (Fragment) OS=Homo sapiens GN=IGLC2 PE=4 SV=1 - [A0A075B6K9_HUMAN]	2	9	73
P04114	Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=2 - [APOB_HUMAN]	81	81	92
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2 - [K22E_HUMAN]	23	34	72
P01861	Ig gamma-4 chain C region OS=Homo sapiens GN=IGHG4 PE=1 SV=1 - [IGHG4_HUMAN]	7	14	73
B9A064	Immunoglobulin lambda-like polypeptide 5 OS=Homo sapiens GN=IGLL5 PE=2 SV=2 - [IGLL5_HUMAN]	3	9	57
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3 - [K1C9_HUMAN]	24	25	68
P21333	Filamin-A OS=Homo sapiens GN=FLNA PE=1 SV=4 - [FLNA_HUMAN]	60	63	76
P02533	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4 - [K1C14_HUMAN]	10	25	56
P08670	Vimentin OS=Homo sapiens GN=VIM PE=1 SV=4 - [VIME_HUMAN]	33	35	50
P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=1 SV=4 - [K1C16_HUMAN]	9	24	53
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3 - [K2C5_HUMAN]	15	29	44
A0A075B6L1	Ig lambda-7 chain C region (Fragment) OS=Homo sapiens GN=IGLC7 PE=1 SV=2 - [A0A075B6L1_HUMAN]	3	7	37
Q562R1	Beta-actin-like protein 2 OS=Homo sapiens GN=ACTBL2 PE=1 SV=2 - [ACTBL_HUMAN]	1	9	56
P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=1 SV=5 - [K2C6B_HUMAN]	2	25	36
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [ALBU_HUMAN]	25	28	42
A0A087WYJ9	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=2 - [A0A087WYJ9_HUMAN]	6	17	35
P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=1 SV=3 - [K2C6A_HUMAN]	2	24	33
P04004	Vitronectin OS=Homo sapiens GN=VTN PE=1 SV=1 - [VTNC_HUMAN]	12	12	29
E7ENL6	Collagen alpha-3(VI) chain OS=Homo sapiens GN=COL6A3 PE=1 SV=2 - [E7ENL6_HUMAN]	36	36	37
P10909	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1 - [CLUS_HUMAN]	14	14	30
P02647	Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1 - [APOA1_HUMAN]	19	19	28
P02751	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=4 - [FINC_HUMAN]	28	28	33
P04220	Ig mu heavy chain disease protein OS=Homo sapiens GN=IGHM PE=1 SV=1 - [MUCB_HUMAN]	1	12	22
P02649	Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1 - [APOE_HUMAN]	18	18	31
P21810	Biglycan OS=Homo sapiens GN=BGN PE=1 SV=2 - [PGS1_HUMAN]	12	12	24
P51888	Prolargin OS=Homo sapiens GN=PRELP PE=1 SV=1 - [PRELP_HUMAN]	14	14	25
P01876	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2 - [IGHA1_HUMAN]	5	13	24
P04406	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 - [G3P_HUMAN]	9	14	21
A0A0B4J1U7	Protein IGHV6-1 (Fragment) OS=Homo sapiens GN=IGHV6-1 PE=1 SV=1 - [A0A0B4J1U7_HUMAN]	5	5	19
A0A0G2JMB2	Ig alpha-2 chain C region (Fragment) OS=Homo sapiens GN=IGHA2 PE=1 SV=1 - [A0A0G2JMB2_HUMAN]	1	9	15
P68871	Hemoglobin subunit beta OS=Homo sapiens GN=HBB PE=1 SV=2 - [HBB_HUMAN]	9	10	20
G1U6W1	Elongation factor 1-alpha OS=Oryctolagus cuniculus PE=3 SV=1 - [G1U6W1_RABIT]	13	13	18
C9JEU5	Fibrinogen gamma chain OS=Homo sapiens GN=FGG PE=1 SV=1 - [C9JEU5_HUMAN]	9	9	12
A0A024R6I7	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=1 - [A0A024R6I7_HUMAN]	12	12	18

P02675	Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2 - [FIBB_HUMAN]	13	13	17
A0A087X0S5	Collagen alpha-1(VI) chain OS=Homo sapiens GN=COL6A1 PE=1 SV=1 - [A0A087X0S5_HUMAN]	13	13	17
K7ERE3	Keratin, type I cytoskeletal 13 OS=Homo sapiens GN=KRT13 PE=1 SV=1 - [K7ERE3_HUMAN]	10	15	19
P14618	Pyruvate kinase PKM OS=Homo sapiens GN=PKM PE=1 SV=4 - [KPYM_HUMAN]	6	16	19
E7ER44	Lactotransferrin OS=Homo sapiens GN=LTF PE=1 SV=1 - [E7ER44_HUMAN]	16	16	16
P02743	Serum amyloid P-component OS=Homo sapiens GN=APCS PE=1 SV=2 - [SAMP_HUMAN]	7	7	12
P06732	Creatine kinase M-type OS=Homo sapiens GN=CKM PE=1 SV=2 - [KCRM_HUMAN]	1	8	14
Q05707	Collagen alpha-1(XIV) chain OS=Homo sapiens GN=COL14A1 PE=1 SV=3 - [COEA1_HUMAN]	13	13	16
E9PKE3	Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1 - [E9PKE3_HUMAN]	4	11	12
P69905	Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=2 - [HBA_HUMAN]	5	5	10
G1S134	Calcium-transporting ATPase OS=Oryctolagus cuniculus GN=ATP2A1 PE=3 SV=2 - [G1S134_RABIT]	1	10	12
P07437	Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV=2 - [TBB5_HUMAN]	2	12	17
G1S110	Tubulin alpha chain OS=Oryctolagus cuniculus GN=TUBA1B PE=3 SV=1 - [G1S110_RABIT]	1	10	16
A0A0C4DGL8	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1 - [A0A0C4DGL8_HUMAN]	6	15	17
Q71U36	Tubulin alpha-1A chain OS=Homo sapiens GN=TUBA1A PE=1 SV=1 - [TBA1A_HUMAN]	1	10	16
P00563	Creatine kinase M-type OS=Oryctolagus cuniculus GN=CKM PE=1 SV=1 - [KCRM_RABIT]	1	8	14
P01024	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2 - [CO3_HUMAN]	16	16	17
P13611	Versican core protein OS=Homo sapiens GN=VCAN PE=1 SV=3 - [CSPG2_HUMAN]	2	11	12
Q08431	Lactadherin OS=Homo sapiens GN=MFG8 PE=1 SV=2 - [MFGM_HUMAN]	2	11	11
Q02413	Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2 - [DSG1_HUMAN]	10	10	12
H3BQN4	Fructose-bisphosphate aldolase OS=Homo sapiens GN=ALDOA PE=1 SV=1 - [H3BQN4_HUMAN]	9	10	13
P51884	Lumican OS=Homo sapiens GN=LUM PE=1 SV=2 - [LUM_HUMAN]	5	9	13
P01023	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=3 - [A2MG_HUMAN]	19	19	22
P68371	Tubulin beta-4B chain OS=Homo sapiens GN=TUBB4B PE=1 SV=1 - [TBB4B_HUMAN]	2	10	15
Q7Z794	Keratin, type II cytoskeletal 1b OS=Homo sapiens GN=KRT77 PE=2 SV=3 - [K2C1B_HUMAN]	8	11	17
P23284	Peptidyl-prolyl cis-trans isomerase B OS=Homo sapiens GN=PP1B PE=1 SV=2 - [PP1B_HUMAN]	12	12	14
P07355	Annexin A2 OS=Homo sapiens GN=ANXA2 PE=1 SV=2 - [ANXA2_HUMAN]	15	15	18
A0A0J9YY99	Uncharacterized protein (Fragment) OS=Homo sapiens PE=1 SV=1 - [A0A0J9YY99_HUMAN]	3	4	10
D6RGZ6	Versican core protein (Fragment) OS=Homo sapiens GN=VCAN PE=1 SV=1 - [D6RGZ6_HUMAN]	1	10	11
O14983	Sarcoplasmic/endoplasmic reticulum calcium ATPase 1 OS=Homo sapiens GN=ATP2A1 PE=1 SV=1 - [AT2A]	2	11	11
P11974	Pyruvate kinase PKM OS=Oryctolagus cuniculus GN=PKM PE=1 SV=4 - [KPYM_RABIT]	1	11	14
P00739	Haptoglobin-related protein OS=Homo sapiens GN=HPR PE=2 SV=2 - [HPTR_HUMAN]	4	13	13
X6R3G6	Lactadherin OS=Homo sapiens GN=MFG8 PE=1 SV=1 - [X6R3G6_HUMAN]	1	10	10
A0A0A0MRZ8	Protein IGKV3D-11 (Fragment) OS=Homo sapiens GN=IGKV3D-11 PE=4 SV=6 - [A0A0A0MRZ8_HUMAN]	2	2	7
A0A0B412B5	Protein IGHV3OR16-9 (Fragment) OS=Homo sapiens GN=IGHV3OR16-9 PE=1 SV=1 - [A0A0B412B5_HUMAN]	1	4	11
P00558	Phosphoglycerate kinase 1 OS=Homo sapiens GN=PGK1 PE=1 SV=3 - [PGK1_HUMAN]	8	8	9
P62805	Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2 - [H4_HUMAN]	6	6	11
P54652	Heat shock-related 70 kDa protein 2 OS=Homo sapiens GN=HSPA2 PE=1 SV=1 - [HSP72_HUMAN]	1	8	9
A0A087WW89	Protein IGHV3-72 OS=Homo sapiens GN=IGHV3-72 PE=1 SV=1 - [A0A087WW89_HUMAN]	1	3	7
A0A0A0MSV6	Complement C1q subcomponent subunit B (Fragment) OS=Homo sapiens GN=C1QB PE=1 SV=6 - [A0A0A0C]	6	6	10
P61626	Lysozyme C OS=Homo sapiens GN=LYZ PE=1 SV=1 - [LYSC_HUMAN]	7	7	8
P01624	Immunoglobulin kappa variable 3-15 OS=Homo sapiens GN=IGKV3-15 PE=1 SV=2 - [KV315_HUMAN]	1	2	7
A0A087X0Q4	Protein IGKV2-40 OS=Homo sapiens GN=IGKV2-40 PE=1 SV=1 - [A0A087X0Q4_HUMAN]	3	4	7

A0A075B6K5	HCG2043239 (Fragment) OS=Homo sapiens GN=IGLV3-9 PE=1 SV=1 - [A0A075B6K5_HUMAN]	3	3	9
P04792	Heat shock protein beta-1 OS=Homo sapiens GN=HSPB1 PE=1 SV=2 - [HSPB1_HUMAN]	5	5	9
Q96PD5	N-acetylmuramoyl-L-alanine amidase OS=Homo sapiens GN=PGLYRP2 PE=1 SV=1 - [PGLYRP2_HUMAN]	5	5	6
G1T6D3	Uncharacterized protein OS=Oryctolagus cuniculus GN=KRT72 PE=3 SV=1 - [G1T6D3_RABIT]	1	5	9
Q86YZ3	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2 - [HORN_HUMAN]	4	4	5
A0A087X0P6	Protein IGKV2D-29 OS=Homo sapiens GN=IGKV2D-29 PE=1 SV=1 - [A0A087X0P6_HUMAN]	2	4	6
P0DMV8	Heat shock 70 kDa protein 1A OS=Homo sapiens GN=HSPA1A PE=1 SV=1 - [HS71A_HUMAN]	7	9	9
GIU9S2	Serum albumin OS=Oryctolagus cuniculus GN=ALB PE=1 SV=1 - [GIU9S2_RABIT]	2	5	6
Q01995	Transgelin OS=Homo sapiens GN=TAGLN PE=1 SV=4 - [TAGL_HUMAN]	5	5	7
P08238	Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4 - [HS90B_HUMAN]	3	6	7
P46406	Glyceraldehyde-3-phosphate dehydrogenase OS=Oryctolagus cuniculus GN=GAPDH PE=1 SV=3 - [G3P_RA]	2	7	10
Q08554	Desmocollin-1 OS=Homo sapiens GN=DSC1 PE=1 SV=2 - [DSC1_HUMAN]	6	6	6
B1ALD9	Periostin OS=Homo sapiens GN=POSTN PE=1 SV=1 - [B1ALD9_HUMAN]	9	9	12
P62937	Peptidyl-prolyl cis-trans isomerase A OS=Homo sapiens GN=PP1A PE=1 SV=2 - [PP1A_HUMAN]	6	6	8
P11021	78 kDa glucose-regulated protein OS=Homo sapiens GN=HSPA5 PE=1 SV=2 - [GRP78_HUMAN]	6	7	7
GI1SP97	Lumican OS=Oryctolagus cuniculus GN=LUM PE=4 SV=1 - [GI1SP97_RABIT]	1	5	6
P06312	Immunoglobulin kappa variable 4-1 OS=Homo sapiens GN=IGKV4-1 PE=1 SV=1 - [KV401_HUMAN]	3	3	7
Q13509	Tubulin beta-3 chain OS=Homo sapiens GN=TUBB3 PE=1 SV=2 - [TBB3_HUMAN]	1	7	10
P60174	Triosephosphate isomerase OS=Homo sapiens GN=TP11 PE=1 SV=3 - [TP1S_HUMAN]	7	7	8
A0A0A0MS10	Peroxisredoxin-1 (Fragment) OS=Homo sapiens GN=PRDX1 PE=1 SV=1 - [A0A0A0MS10_HUMAN]	4	5	6
P06576	ATP synthase subunit beta, mitochondrial OS=Homo sapiens GN=ATP5B PE=1 SV=3 - [ATPB_HUMAN]	8	8	8
GI1SKT4	ATP synthase subunit alpha OS=Oryctolagus cuniculus GN=ATP5A1 PE=3 SV=1 - [GI1SKT4_RABIT]	7	7	7
P01619	Immunoglobulin kappa variable 3-20 OS=Homo sapiens GN=IGKV3-20 PE=1 SV=2 - [KV320_HUMAN]	2	4	5
P13929	Beta-enolase OS=Homo sapiens GN=ENO3 PE=1 SV=5 - [ENOB_HUMAN]	8	9	9
A0A075B6H7	Protein IGKV3-7 (Fragment) OS=Homo sapiens GN=IGKV3-7 PE=4 SV=1 - [A0A075B6H7_HUMAN]	1	2	7
A0A0G2JL54	Complement C4-B OS=Homo sapiens GN=C4B_2 PE=1 SV=1 - [A0A0G2JL54_HUMAN]	5	5	5
P06331	Immunoglobulin heavy variable 4-34 OS=Homo sapiens GN=IGHV4-34 PE=1 SV=2 - [HV434_HUMAN]	3	4	5
P16615	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 OS=Homo sapiens GN=ATP2A2 PE=1 SV=1 - [AT2A]	1	5	6
P02747	Complement C1q subcomponent subunit C OS=Homo sapiens GN=C1QC PE=1 SV=3 - [C1QC_HUMAN]	6	6	6
P04432	Immunoglobulin kappa variable 1D-39 OS=Homo sapiens GN=IGKV1D-39 PE=3 SV=2 - [KYD39_HUMAN]	2	3	7
Q5T749	Keratinocyte proline-rich protein OS=Homo sapiens GN=KPRP PE=1 SV=1 - [KPRP_HUMAN]	6	6	6
P04003	C4b-binding protein alpha chain OS=Homo sapiens GN=C4BPA PE=1 SV=2 - [C4BPA_HUMAN]	5	5	6
P01709	Immunoglobulin lambda variable 2-8 OS=Homo sapiens GN=IGLV2-8 PE=1 SV=2 - [LV208_HUMAN]	2	2	5
A0A1B0GW44	Cathepsin D OS=Homo sapiens GN=CTSD PE=1 SV=1 - [A0A1B0GW44_HUMAN]	5	5	6
P02671	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2 - [FIBA_HUMAN]	5	6	6
A0A0C4DH67	Protein IGKV1-8 (Fragment) OS=Homo sapiens GN=IGKV1-8 PE=1 SV=1 - [A0A0C4DH67_HUMAN]	2	3	6
P12110	Collagen alpha-2(VI) chain OS=Homo sapiens GN=COL6A2 PE=1 SV=4 - [CO6A2_HUMAN]	6	6	6
P01700	Immunoglobulin lambda variable 1-47 OS=Homo sapiens GN=IGLV1-47 PE=1 SV=2 - [LV147_HUMAN]	3	3	4
P08519	Apolipoprotein(a) OS=Homo sapiens GN=LPA PE=1 SV=1 - [APOA_HUMAN]	6	6	7
P63244	Receptor of activated protein C kinase 1 OS=Homo sapiens GN=RACK1 PE=1 SV=3 - [RACK1_HUMAN]	4	4	5
GI17F1	Histone H2B OS=Oryctolagus cuniculus GN=LOC100348835 PE=3 SV=1 - [GI17F1_RABIT]	4	4	5
F5H6Q2	Polyubiquitin-C (Fragment) OS=Homo sapiens GN=UBC PE=1 SV=8 - [F5H6Q2_HUMAN]	3	3	4
A0A075B7B8	Protein IGHV3OR16-12 (Fragment) OS=Homo sapiens GN=IGHV3OR16-12 PE=1 SV=1 - [A0A075B7B8_HU]	1	2	4

P06727	Apolipoprotein A-IV OS=Homo sapiens GN=APOA4 PE=1 SV=3 - [APOA4_HUMAN]	6	6	6	6
Q08380	Galectin-3-binding protein OS=Homo sapiens GN=LGALS3BP PE=1 SV=1 - [LG3BP_HUMAN]	7	7	7	8
P04083	Annexin A1 OS=Homo sapiens GN=ANXA1 PE=1 SV=2 - [ANXA1_HUMAN]	7	7	7	7
P07737	Profilin-1 OS=Homo sapiens GN=PFN1 PE=1 SV=2 - [PROF1_HUMAN]	5	5	5	5
P07951	Tropomyosin beta chain OS=Homo sapiens GN=TPM2 PE=1 SV=1 - [TPM2_HUMAN]	1	1	1	5
G1T650	Tropomyosin alpha-1 chain OS=Oryctolagus cuniculus GN=TPM1 PE=3 SV=2 - [G1T650_RABIT]	1	1	1	5
Q9Y6R7	IgGfc-binding protein OS=Homo sapiens GN=FCGBP PE=1 SV=3 - [FCGBP_HUMAN]	6	6	6	6
P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens GN=KRT4 PE=1 SV=4 - [K2C4_HUMAN]	4	4	4	7
P01031	Complement C5 OS=Homo sapiens GN=C5 PE=1 SV=4 - [C05_HUMAN]	10	10	10	10
P01599	Immunoglobulin kappa variable 1-17 OS=Homo sapiens GN=IGKV1-17 PE=1 SV=2 - [KV117_HUMAN]	3	3	3	5
Q00610	Clathrin heavy chain 1 OS=Homo sapiens GN=CLTC PE=1 SV=5 - [CLH1_HUMAN]	7	7	7	8
G1TBC1	Endoplasmin OS=Oryctolagus cuniculus GN=HSP90B1 PE=3 SV=1 - [G1TBC1_RABBIT]	5	5	5	6
Q16555	Dihydropyrimidinase-related protein 2 OS=Homo sapiens GN=DPYSL2 PE=1 SV=1 - [DPYL2_HUMAN]	5	5	5	6
P07360	Complement component C8 gamma chain OS=Homo sapiens GN=C8G PE=1 SV=3 - [C08G_HUMAN]	5	5	5	5
P01718	Immunoglobulin lambda variable 3-27 OS=Homo sapiens GN=IGLV3-27 PE=1 SV=2 - [LV327_HUMAN]	3	3	3	4
P07996	Thrombospondin-1 OS=Homo sapiens GN=THBS1 PE=1 SV=2 - [TSP1_HUMAN]	3	3	3	5
A0A087WV45	Transferrin OS=Homo sapiens GN=TTR PE=1 SV=1 - [A0A087WV45_HUMAN]	5	5	5	5
P07093	Glia-derived nexin OS=Homo sapiens GN=SERPINE2 PE=1 SV=1 - [GDN_HUMAN]	4	4	4	5
Q8N1N4	Keratin, type II cytoskeletal 78 OS=Homo sapiens GN=KRT78 PE=2 SV=2 - [K2C78_HUMAN]	4	4	4	5
E9PHK0	Tetranectin OS=Homo sapiens GN=CLEC3B PE=1 SV=1 - [E9PHK0_HUMAN]	4	4	4	4
Q5T985	Inter-alpha-trypsin inhibitor heavy chain H2 OS=Homo sapiens GN=ITIH2 PE=1 SV=1 - [Q5T985_HUMAN]	7	7	7	7
A0A075B6K4	HCG2043238 (Fragment) OS=Homo sapiens GN=IGLV3-10 PE=1 SV=1 - [A0A075B6K4_HUMAN]	3	3	3	4
A0A0C4DH31	Protein IGHV1-18 (Fragment) OS=Homo sapiens GN=IGHV1-18 PE=1 SV=1 - [A0A0C4DH31_HUMAN]	2	2	2	3
A0A0C4DH25	Immunoglobulin kappa variable 3-20 OS=Homo sapiens GN=IGKV3-20 PE=3 SV=1 - [KVD20_HUMAN]	1	1	1	4
G1SE61	Filamin-B OS=Oryctolagus cuniculus GN=FLNB PE=4 SV=1 - [G1SE61_RABBIT]	1	1	1	6
E7END7	Ras-related protein Rab-1A OS=Homo sapiens GN=RAB1A PE=1 SV=1 - [E7END7_HUMAN]	5	5	5	6
F5H018	GTP-binding nuclear protein Ran (Fragment) OS=Homo sapiens GN=GAN PE=1 SV=8 - [F5H018_HUMAN]	4	4	4	4
P06733	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [ENOA_HUMAN]	4	4	4	6
P27169	Serum paraoxonase/arylesterase 1 OS=Homo sapiens GN=PON1 PE=1 SV=3 - [PON1_HUMAN]	5	5	5	5
G1T2J0	Uncharacterized protein OS=Oryctolagus cuniculus GN=JUP PE=4 SV=1 - [G1T2J0_RABIT]	8	8	8	8
A0A075B6R2	Protein IGHV4-4 (Fragment) OS=Homo sapiens GN=IGHV4-4 PE=1 SV=2 - [A0A075B6R2_HUMAN]	1	1	1	3
A0A0A0M551	Gelsolin OS=Homo sapiens GN=GSN PE=1 SV=1 - [A0A0A0M551_HUMAN]	5	5	5	5
P02099	Hemoglobin subunit gamma OS=Oryctolagus cuniculus GN=HGB PE=2 SV=2 - [HBG_RABIT]	1	1	1	4
P35542	Serum amyloid A-4 protein OS=Homo sapiens GN=SAA4 PE=1 SV=2 - [SAA4_HUMAN]	5	5	5	9
P01701	Immunoglobulin lambda variable 1-51 OS=Homo sapiens GN=IGLV1-51 PE=1 SV=2 - [LV151_HUMAN]	3	3	3	5
C9J9K3	40S ribosomal protein SA (Fragment) OS=Homo sapiens GN=RP5A PE=1 SV=8 - [C9J9K3_HUMAN]	4	4	4	4
P01782	Immunoglobulin heavy variable 3-9 OS=Homo sapiens GN=IGHV3-9 PE=1 SV=2 - [HV309_HUMAN]	2	2	2	7
P02792	Ferritin light chain OS=Homo sapiens GN=FTL PE=1 SV=2 - [FRIL_HUMAN]	3	3	3	3
P01602	Immunoglobulin kappa variable 1-5 OS=Homo sapiens GN=IGKV1-5 PE=1 SV=2 - [KV105_HUMAN]	2	2	2	4
A0A0G2JRK6	Uncharacterized protein (Fragment) OS=Homo sapiens GN=IGKV1-5 PE=1 SV=4 - [A0A0G2JRK6_HUMAN]	2	2	2	5
A0A075B7D0	Protein IGHV1OR15-1 (Fragment) OS=Homo sapiens GN=IGHV1OR15-1 PE=1 SV=1 - [A0A075B7D0_HUMAN]	2	2	2	4
D6RBL5	Annexin OS=Homo sapiens GN=ANXA5 PE=1 SV=1 - [D6RBL5_HUMAN]	4	4	4	4
A0A0B4J2D9	Protein IGKV1D-13 (Fragment) OS=Homo sapiens GN=IGKV1D-13 PE=1 SV=1 - [A0A0B4J2D9_HUMAN]	2	2	2	4

P60981	Destrin OS=Homo sapiens GN=DSTN PE=1 SV=3 - [DEST_HUMAN]	4	4	4
P48047	ATP synthase subunit O, mitochondrial OS=Homo sapiens GN=ATP5O PE=1 SV=1 - [ATPO_HUMAN]	3	3	3
P01780	Immunoglobulin heavy variable 3-7 OS=Homo sapiens GN=IGHV3-7 PE=1 SV=2 - [HV307_HUMAN]	3	5	8
Q3BDU5	Prelamin-A/C OS=Homo sapiens GN=LMNA PE=1 SV=1 - [Q3BDU5_HUMAN]	4	4	4
G3V1A4	Cofilin 1 (Non-muscle), isoform CRA_a OS=Homo sapiens GN=CFL1 PE=1 SV=1 - [G3V1A4_HUMAN]	3	3	3
H3BMH2	Ras-related protein Rab-11A (Fragment) OS=Homo sapiens GN=RAB11A PE=4 SV=1 - [H3BMH2_HUMAN]	4	4	4
D6RD17	Immunoglobulin J chain (Fragment) OS=Homo sapiens GN=JCHAIN PE=1 SV=8 - [D6RD17_HUMAN]	3	3	3
P63104	14-3-3 protein zeta/delta OS=Homo sapiens GN=YWHAZ PE=1 SV=1 - [1433Z_HUMAN]	4	6	6
P05090	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1 - [APOD_HUMAN]	4	4	4
P01714	Immunoglobulin lambda variable 3-19 OS=Homo sapiens GN=IGLV3-19 PE=1 SV=2 - [LV319_HUMAN]	2	2	3
P23083	Immunoglobulin heavy variable 1-2 OS=Homo sapiens GN=IGHV1-2 PE=1 SV=2 - [HV102_HUMAN]	1	2	2
Q9Y490	Talin-1 OS=Homo sapiens GN=TLN1 PE=1 SV=3 - [TLN1_HUMAN]	5	5	5
A0A075B6R9	Protein IGKV2D-24 (Fragment) OS=Homo sapiens GN=IGKV2D-24 PE=4 SV=1 - [A0A075B6R9_HUMAN]	2	3	3
Q14195	Dihydropyrimidinase-related protein 3 OS=Homo sapiens GN=DPYSL3 PE=1 SV=1 - [DPYL3_HUMAN]	5	6	6
P07858	Cathepsin B OS=Homo sapiens GN=CTSB PE=1 SV=3 - [CATB_HUMAN]	3	3	3
E9PC84	Tenascin OS=Homo sapiens GN=TNC PE=1 SV=1 - [E9PC84_HUMAN]	3	3	3
P07900	Heat shock protein HSP 90-alpha OS=Homo sapiens GN=HSP90AA1 PE=1 SV=5 - [HS90A_HUMAN]	1	2	2
Q6NZ12	Polymerase I and transcript release factor OS=Homo sapiens GN=PTRF PE=1 SV=1 - [PTRF_HUMAN]	3	3	3
U3KPH3	Uncharacterized protein OS=Oryctolagus cuniculus GN=SLC25A5 PE=3 SV=1 - [U3KPH3_RABIT]	4	4	4
P55058	Phospholipid transfer protein OS=Homo sapiens GN=PLTP PE=1 SV=1 - [PLTP_HUMAN]	4	4	4
C9JG13	Thymidine phosphorylase (Fragment) OS=Homo sapiens GN=TYMP PE=1 SV=1 - [C9JG13_HUMAN]	3	3	3
G1SQX9	Uncharacterized protein OS=Oryctolagus cuniculus GN=SFN PE=3 SV=1 - [G1SQX9_RABIT]	3	6	6
P04430	Immunoglobulin kappa variable 1-16 OS=Homo sapiens GN=IGKV1-16 PE=1 SV=2 - [KV116_HUMAN]	2	3	4
G1T652	Fructose-bisphosphate aldolase OS=Oryctolagus cuniculus GN=ALDOC PE=3 SV=1 - [G1T652_RABIT]	1	2	3
P01011	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2 - [AACT_HUMAN]	6	6	6
P02749	Beta-2-glycoprotein 1 OS=Homo sapiens GN=APOH PE=1 SV=3 - [APOH_HUMAN]	4	4	5
G1U4G9	Chloride intracellular channel protein OS=Oryctolagus cuniculus GN=CLIC1 PE=3 SV=1 - [G1U4G9_RABIT]	3	3	3
P81605	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 - [DCD_HUMAN]	2	2	3
E5RK99	Collagen triple helix repeat-containing protein 1 OS=Homo sapiens GN=CTHRC1 PE=1 SV=1 - [E5RK99_HU]	3	3	3
P29508	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2 - [SPB3_HUMAN]	3	3	3
P12273	Prolactin-inducible protein OS=Homo sapiens GN=PIP PE=1 SV=1 - [PIP_HUMAN]	3	3	3
C9J592	Ras-related protein Rab-7a (Fragment) OS=Homo sapiens GN=RAB7A PE=1 SV=1 - [C9J592_HUMAN]	4	4	4
P51148	Ras-related protein Rab-5C OS=Homo sapiens GN=RAB5C PE=1 SV=2 - [RAB5C_HUMAN]	3	3	3
A0A0C4DH38	Protein IGHV5-51 (Fragment) OS=Homo sapiens GN=IGHV5-51 PE=1 SV=1 - [A0A0C4DH38_HUMAN]	1	3	3
A0A0J9YXX1	Uncharacterized protein (Fragment) OS=Homo sapiens GN=PE=1 SV=1 - [A0A0J9YXX1_HUMAN]	1	3	3
P69891	Hemoglobin subunit gamma-1 OS=Homo sapiens GN=HBG1 PE=1 SV=2 - [HBG1_HUMAN]	1	2	4
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN=ITIH4 PE=1 SV=4 - [ITIH4_HUMAN]	4	4	5
A0A0B4J1R6	Transketolase OS=Homo sapiens GN=TKT PE=1 SV=1 - [A0A0B4J1R6_HUMAN]	3	3	3
G1T5H0	Uncharacterized protein OS=Oryctolagus cuniculus GN=HTRA1 PE=4 SV=2 - [G1T5H0_RABIT]	3	3	3
A8MW50	L-lactate dehydrogenase (Fragment) OS=Homo sapiens GN=LDHB PE=1 SV=1 - [A8MW50_HUMAN]	2	3	3
H9KV48	Plasma protease C1 inhibitor OS=Homo sapiens GN=SERPING1 PE=1 SV=1 - [H9KV48_HUMAN]	4	4	4
P01042	Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2 - [KNG1_HUMAN]	6	6	6
P00338	L-lactate dehydrogenase A chain OS=Homo sapiens GN=LDHA PE=1 SV=2 - [LDHA_HUMAN]	2	3	3

A0A0B411Y8	Protein IGLV9-49 OS=Homo sapiens GN=IGLV9-49 PE=1 SV=1 - [A0A0B411Y8_HUMAN]	2	2	2
P01593	Immunoglobulin kappa variable ID-33 OS=Homo sapiens GN=IGKV1D-33 PE=1 SV=2 - [KYD33_HUMAN]	1	1	3
G1TG00	Uncharacterized protein OS=Oryctolagus cuniculus GN=RPS3 PE=4 SV=1 - [G1TG00_RABIT]	4	4	4
A0A0C4DH24	Protein IGKV6-21 (Fragment) OS=Homo sapiens GN=IGKV6-21 PE=1 SV=1 - [A0A0C4DH24_HUMAN]	1	3	4
S4R3Y4	Protein AMBP OS=Homo sapiens GN=AMBP PE=1 SV=1 - [S4R3Y4_HUMAN]	2	2	3
A0A0C4DH34	Protein IGHV4-28 (Fragment) OS=Homo sapiens GN=IGHV4-28 PE=1 SV=1 - [A0A0C4DH34_HUMAN]	3	3	3
A0A075B6I9	Protein IGLV7-46 (Fragment) OS=Homo sapiens GN=IGLV7-46 PE=1 SV=4 - [A0A075B6I9_HUMAN]	2	2	2
A0A0A0MT36	Protein IGKV6D-21 (Fragment) OS=Homo sapiens GN=IGKV6D-21 PE=4 SV=1 - [A0A0A0MT36_HUMAN]	1	3	5
P13639	Elongation factor 2 OS=Homo sapiens GN=EF2 PE=1 SV=4 - [EF2_HUMAN]	4	4	4
P30101	Protein disulfide-isomerase A3 OS=Homo sapiens GN=PDIA3 PE=1 SV=4 - [PDIA3_HUMAN]	4	4	4
P35442	Thrombospondin-2 OS=Homo sapiens GN=THBS2 PE=1 SV=2 - [TSP2_HUMAN]	4	5	5
I3L1H9	Zymogen granule protein 16 homolog B (Fragment) OS=Homo sapiens GN=ZG16B PE=1 SV=1 - [I3L1H9_HUMAN]	2	2	2
P01601	Immunoglobulin kappa variable ID-16 (Fragment) OS=Homo sapiens GN=IGKV1D-16 PE=3 SV=2 - [KVD16_HUMAN]	1	2	3
P01687	Ig kappa chain V region BS-5 OS=Oryctolagus cuniculus PE=1 SV=1 - [KV06_RABIT]	2	2	2
P05109	Protein S100-A8 OS=Homo sapiens GN=S100A8 PE=1 SV=1 - [S100A8_HUMAN]	3	3	3
G1SPPO	Uncharacterized protein OS=Oryctolagus cuniculus GN=STOM PE=4 SV=1 - [G1SPPO_RABIT]	2	2	2
G1TL92	Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=2 - [G1TL92_RABIT]	3	3	3
P31946	14-3-3 protein beta/alpha OS=Homo sapiens GN=YWHAB PE=1 SV=3 - [1433B_HUMAN]	1	4	4
C9J1Z8	ADP-ribosylation factor 5 (Fragment) OS=Homo sapiens GN=ARF5 PE=1 SV=1 - [C9J1Z8_HUMAN]	2	2	2
G1STF7	Serotransferrin OS=Oryctolagus cuniculus GN=TF PE=3 SV=1 - [G1STF7_RABIT]	2	2	2
P31025	Lipocalin-1 OS=Homo sapiens GN=LCN1 PE=1 SV=1 - [LCN1_HUMAN]	3	3	3
Q8IV08	Phospholipase D3 OS=Homo sapiens GN=PLD3 PE=1 SV=1 - [PLD3_HUMAN]	3	3	3
P51911	Calponin-1 OS=Homo sapiens GN=CNN1 PE=1 SV=2 - [CNN1_HUMAN]	2	2	2
Q5TBF5	Mimcan (Fragment) OS=Homo sapiens GN=OGN PE=1 SV=1 - [Q5TBF5_HUMAN]	2	2	2
F8VFP3	Myosin light polypeptide 6 (Fragment) OS=Homo sapiens GN=MYL6 PE=1 SV=1 - [F8VFP3_HUMAN]	2	2	3
G1TUX2	Aconitate hydratase, mitochondrial OS=Oryctolagus cuniculus GN=ACO2 PE=3 SV=2 - [G1TUX2_RABIT]	2	2	2
B1AH77	Ras-related C3 botulinum toxin substrate 2 OS=Homo sapiens GN=RAC2 PE=1 SV=1 - [B1AH77_HUMAN]	4	4	4
P59665	Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1 - [DEF1_HUMAN]	3	3	3
E9PNW4	CD59 glycoprotein OS=Homo sapiens GN=CD59 PE=1 SV=1 - [E9PNW4_HUMAN]	3	3	3
P15924	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3 - [DESP_HUMAN]	5	5	5
P00747	Plasminogen OS=Homo sapiens GN=PLG PE=1 SV=2 - [PLMN_HUMAN]	3	3	4
F8VPE8	60S acidic ribosomal protein P0 (Fragment) OS=Homo sapiens GN=RPLP0 PE=1 SV=1 - [F8VPE8_HUMAN]	3	3	3
Q8LUX7	Adipocyte enhancer-binding protein 1 OS=Homo sapiens GN=AEBP1 PE=1 SV=1 - [AEBP1_HUMAN]	3	3	3
U3KMU3	Alpha-1,4 glucan phosphorylase OS=Oryctolagus cuniculus GN=PYGM PE=3 SV=1 - [U3KMU3_RABIT]	2	4	4
E9PND2	Cysteine and glycine-rich protein 1 (Fragment) OS=Homo sapiens GN=CSRP1 PE=1 SV=1 - [E9PND2_HUMAN]	1	1	2
X6RJP6	Transgelin-2 (Fragment) OS=Homo sapiens GN=TAGLN2 PE=1 SV=1 - [X6RJP6_HUMAN]	6	6	8
E9PQ63	Carbonyl reductase [NADPH] 1 OS=Homo sapiens GN=CBR1 PE=1 SV=1 - [E9PQ63_HUMAN]	2	2	2
G3XAL0	Malate dehydrogenase OS=Homo sapiens GN=MDH2 PE=1 SV=1 - [G3XAL0_HUMAN]	2	2	2
P19827	Inter-alpha-trypsin inhibitor heavy chain H1 OS=Homo sapiens GN=ITIH1 PE=1 SV=3 - [ITIH1_HUMAN]	4	4	4
G1TRG8	Uncharacterized protein OS=Oryctolagus cuniculus GN=GNAI2 PE=4 SV=1 - [G1TRG8_RABIT]	3	3	3
F8VQR7	Cysteine and glycine-rich protein 2 OS=Homo sapiens GN=CSRP2 PE=1 SV=1 - [F8VQR7_HUMAN]	3	3	3
Q5H9A7	Metalloproteinase inhibitor 1 OS=Homo sapiens GN=TIMP1 PE=1 SV=1 - [Q5H9A7_HUMAN]	2	2	2
A0A075B6I0	Protein IGLV8-61 (Fragment) OS=Homo sapiens GN=IGLV8-61 PE=1 SV=7 - [A0A075B6I0_HUMAN]	2	2	3

C9JC71	Low affinity immunoglobulin gamma Fc region receptor III-A (Fragment) OS=Homo sapiens GN=FCGR3A P	2	2	2
U3KMP9	Uncharacterized protein OS=Oryctolagus cuniculus GN=LOC100341294 PE=3 SV=1 - [U3KMP9_RABIT]	2	2	2
P02748	Complement component C9 OS=Homo sapiens GN=C9 PE=1 SV=2 - [CO9_HUMAN]	5	5	5
Q99714	3-hydroxyacyl-CoA dehydrogenase type-2 OS=Homo sapiens GN=HSD17B10 PE=1 SV=3 - [HCD2_HUMAN]	2	2	2
E7ESV4	Ras-related protein Rap-1b (Fragment) OS=Homo sapiens GN=RAP1B PE=1 SV=1 - [E7ESV4_HUMAN]	2	2	3
P39019	40S ribosomal protein S19 OS=Homo sapiens GN=RPS19 PE=1 SV=2 - [RS19_HUMAN]	2	2	2
Q06828	Fibromodulin OS=Homo sapiens GN=FMOD PE=1 SV=2 - [FMOD_HUMAN]	3	3	3
G1TWA7	Fibrinogen alpha chain OS=Oryctolagus cuniculus GN=FGA PE=4 SV=2 - [GITWA7_RABIT]	1	2	2
P32119	Peroxiredoxin-2 OS=Homo sapiens GN=PRDX2 PE=1 SV=5 - [PRDX2_HUMAN]	1	2	2
Q6KB66	Keratin, type II cytoskeletal 80 OS=Homo sapiens GN=KR180 PE=1 SV=2 - [K2C80_HUMAN]	3	3	3
F8VXZ8	Decorin (Fragment) OS=Homo sapiens GN=DCN PE=1 SV=1 - [F8VXZ8_HUMAN]	3	3	3
P48745	Protein NOV homolog OS=Homo sapiens GN=NOV PE=1 SV=1 - [NOV_HUMAN]	5	5	5
A0A0A0MR02	Voltage-dependent anion-selective channel protein 2 (Fragment) OS=Homo sapiens GN=VDAC2 PE=1 SV=	2	2	2
P09382	Galectin-1 OS=Homo sapiens GN=LGAL51 PE=1 SV=2 - [LEG1_HUMAN]	2	2	2
P26038	Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 - [MOES_HUMAN]	7	7	7
Q6PID9	Aggrecan OS=Homo sapiens GN=ACAN PE=1 SV=1 - [Q6PID9_HUMAN]	3	3	3
P05089	Arginase-1 OS=Homo sapiens GN=ARG1 PE=1 SV=2 - [ARG1_HUMAN]	2	2	2
G1U025	Myosin light chain 1/3, skeletal muscle isoform OS=Oryctolagus cuniculus GN=MYL1 PE=4 SV=1 - [G1U025	3	4	4
C9JH87	Voltage-dependent anion-selective channel protein 1 (Fragment) OS=Homo sapiens GN=VDAC1 PE=1 SV=	3	3	3
Q5VVC9	60S ribosomal protein L11 (Fragment) OS=Homo sapiens GN=RPL11 PE=1 SV=1 - [Q5VVC9_HUMAN]	2	2	2
Q14956	Transmembrane glycoprotein NMB OS=Homo sapiens GN=GPNMB PE=1 SV=2 - [GPNMB_HUMAN]	1	1	2
G3V513	Alpha-1-antichymotrypsin (Fragment) OS=Homo sapiens GN=SERPINA3 PE=4 SV=2 - [G3V513_HUMAN]	2	2	2
K7ESA3	Synaptic vesicle membrane protein VAT-1 homolog (Fragment) OS=Homo sapiens GN=VAT1 PE=1 SV=1 -	2	2	2
Q9TT13	Voltage-dependent anion-selective channel protein 3 OS=Oryctolagus cuniculus GN=VDAC3 PE=2 SV=1 - [1	2	2
G1U0D0	Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=1 - [G1U0D0_RABIT]	1	1	2
G1SZ15	Uncharacterized protein OS=Oryctolagus cuniculus PE=1 SV=1 - [G1SZ15_RABIT]	3	3	3
Q16853	Membrane primary amine oxidase OS=Homo sapiens GN=AOC3 PE=1 SV=3 - [AOC3_HUMAN]	3	3	3
P08246	Neutrophil elastase OS=Homo sapiens GN=ELANE PE=1 SV=1 - [ELNE_HUMAN]	2	2	2
G1U4H8	Uncharacterized protein OS=Oryctolagus cuniculus GN=ACTN3 PE=4 SV=1 - [G1U4H8_RABIT]	1	3	3
A0A075B6S9	Protein IGKV1-37 (Fragment) OS=Homo sapiens GN=IGKV1D-37 PE=4 SV=7 - [A0A075B6S9_HUMAN]	1	2	3
P05546	Heparin cofactor 2 OS=Homo sapiens GN=SERPIND1 PE=1 SV=3 - [HEP2_HUMAN]	3	3	3
Q15365	Poly(rC)-binding protein 1 OS=Homo sapiens GN=PCBP1 PE=1 SV=2 - [PCBP1_HUMAN]	4	4	4
Q9NZN4	EH domain-containing protein 2 OS=Homo sapiens GN=EHD2 PE=1 SV=2 - [EHD2_HUMAN]	2	2	2
G1T1V5	Uncharacterized protein OS=Oryctolagus cuniculus GN=CA3 PE=4 SV=1 - [G1T1V5_RABIT]	1	1	1
H7C106	Reticulon (Fragment) OS=Homo sapiens GN=RTN4 PE=1 SV=1 - [H7C106_HUMAN]	1	1	1
A0A0B4J2H0	Protein IGHV1-69-2 (Fragment) OS=Homo sapiens GN=IGHV1-69-2 PE=1 SV=1 - [A0A0B4J2H0_HUMAN]	1	1	1
H0Y8L3	Transforming growth factor-beta-induced protein ig-h3 (Fragment) OS=Homo sapiens GN=TFBFI PE=1 SV	3	3	3
P46940	Ras GTPase-activating-like protein IQGAP1 OS=Homo sapiens GN=IQGAP1 PE=1 SV=1 - [IQGA1_HUMAN]	5	5	5
O14773	Tripeptidyl-peptidase 1 OS=Homo sapiens GN=TPP1 PE=1 SV=2 - [TPP1_HUMAN]	2	2	2
Q9UGM3	Deleted in malignant brain tumors 1 protein OS=Homo sapiens GN=DMBT1 PE=1 SV=2 - [DMBT1_HUMAN]	1	1	1
P01703	Immunoglobulin lambda variable 1-40 OS=Homo sapiens GN=IGLV1-40 PE=1 SV=2 - [LV140_HUMAN]	2	2	2
P47845	Galectin-3 OS=Oryctolagus cuniculus GN=LGALS3 PE=2 SV=2 - [LEG3_RABIT]	1	1	1
P63167	Dynein light chain 1, cytoplasmic OS=Homo sapiens GN=DYNLL1 PE=1 SV=1 - [DYL1_HUMAN]	1	1	1

HOYKF0	Electron transfer flavoprotein subunit alpha, mitochondrial (Fragment) OS=Homo sapiens GN=ETFA PE=1	2	2	2
P29972	Aquaporin-1 OS=Homo sapiens GN=AQP1 PE=1 SV=3 - [AQP1_HUMAN]	1	1	1
HOYHD8	Cysteine-rich protein 2 (Fragment) OS=Homo sapiens GN=CRIP2 PE=1 SV=1 - [HOYHD8_HUMAN]	3	3	3
Q08188	Protein-glutamine gamma-glutamyltransferase E OS=Homo sapiens GN=TGM3 PE=1 SV=4 - [TGM3_HUMA]	4	4	4
P27797	Cairreticulin OS=Homo sapiens GN=CALR PE=1 SV=1 - [CALR_HUMAN]	1	1	1
P01766	Immunoglobulin heavy variable 3-13 OS=Homo sapiens GN=IGHV3-13 PE=1 SV=2 - [HV313_HUMAN]	1	2	3
G1TV31	Actin-related protein 2/3 complex subunit 5 OS=Oryctolagus cuniculus GN=ARPC5 PE=3 SV=1 - [G1TV31_	1	1	1
A0A0B411V0	Protein IGHV3-15 (Fragment) OS=Homo sapiens GN=IGHV3-15 PE=1 SV=1 - [A0A0B411V0_HUMAN]	1	2	2
P02462	Collagen alpha-1(IV) chain OS=Homo sapiens GN=COL4A1 PE=1 SV=3 - [CO4A1_HUMAN]	1	1	1
G1TFA7	Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=1 - [G1TFA7_RABIT]	2	2	2
G1T7D1	Uncharacterized protein OS=Oryctolagus cuniculus GN=HIST1H1C PE=3 SV=1 - [G1T7D1_RABIT]	2	2	2
C9JEV0	Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE=1 SV=1 - [C9JEV0_HUMAN]	1	1	1
F5H7Z2	Microfibrillar-associated protein 5 OS=Homo sapiens GN=MFAP5 PE=1 SV=1 - [F5H7Z2_HUMAN]	1	1	1
C9J0D1	Histone H2A OS=Homo sapiens GN=H2AFV PE=1 SV=1 - [C9J0D1_HUMAN]	1	1	1
D6RHZ6	Alcohol dehydrogenase 1B OS=Homo sapiens GN=ADH1B PE=1 SV=2 - [D6RHZ6_HUMAN]	3	3	3
C9K0Z8	Nucleoside diphosphate kinase A (Fragment) OS=Homo sapiens GN=NME1 PE=1 SV=1 - [C9K0Z8_HUMAN]	1	1	1
A0A0C4DH36	Protein IGHV3-38 (Fragment) OS=Homo sapiens GN=IGHV3-38 PE=1 SV=1 - [A0A0C4DH36_HUMAN]	1	1	1
U3KPF0	Uncharacterized protein OS=Oryctolagus cuniculus GN=CPNE3 PE=4 SV=1 - [U3KPF0_RABIT]	1	1	1
A0A0U1RQ79	Synaptophysin-like protein 1 (Fragment) OS=Homo sapiens GN=SYPL1 PE=1 SV=1 - [A0A0U1RQ79_HUMA]	1	1	1
Q14574	Desmocollin-3 OS=Homo sapiens GN=DSC3 PE=1 SV=3 - [DSC3_HUMAN]	1	1	1
Q13835	Plakophilin-1 OS=Homo sapiens GN=PKP1 PE=1 SV=2 - [PKP1_HUMAN]	1	1	1
P30050	60S ribosomal protein L12 OS=Homo sapiens GN=RPL12 PE=1 SV=1 - [RL12_HUMAN]	2	2	2
A0A087WUQ6	Glutathione peroxidase OS=Homo sapiens GN=GPX1 PE=1 SV=1 - [A0A087WUQ6_HUMAN]	1	1	1
F8VQQ8	Ubiquitin-conjugating enzyme E2 N OS=Homo sapiens GN=UBE2N PE=1 SV=1 - [F8VQQ8_HUMAN]	1	1	1
A0A0G2JS06	Uncharacterized protein (Fragment) OS=Homo sapiens PE=4 SV=1 - [A0A0G2JS06_HUMAN]	1	1	1
J3KSS0	60S ribosomal protein L26 OS=Homo sapiens GN=RPL26 PE=4 SV=1 - [J3KSS0_HUMAN]	1	1	1
P55001	Microfibrillar-associated protein 2 OS=Homo sapiens GN=MFAP2 PE=2 SV=1 - [MFAP2_HUMAN]	1	1	1
HOY9X3	Programmed cell death protein 6 (Fragment) OS=Homo sapiens GN=PDCD6 PE=1 SV=1 - [HOY9X3_HUMA]	1	1	1
A0A140T958	HLA class II histocompatibility antigen, DRB1-16 beta chain (Fragment) OS=Homo sapiens GN=HLA-DRB1	1	1	1
F5H1V0	Complement C1r subcomponent OS=Homo sapiens GN=C1R PE=1 SV=1 - [F5H1V0_HUMAN]	1	1	1
E9PLT1	Platelet glycoprotein 4 OS=Homo sapiens GN=CD36 PE=1 SV=1 - [E9PLT1_HUMAN]	1	1	1
Q09666	Neuroblast differentiation-associated protein AHNAK OS=Homo sapiens GN=AHNAK PE=1 SV=2 - [AHNK_H	1	1	1
J3QRJ3	Thy-1 membrane glycoprotein OS=Homo sapiens GN=THY1 PE=1 SV=1 - [J3QRJ3_HUMAN]	1	1	1
A6NNI4	Tetraspanin OS=Homo sapiens GN=CD9 PE=1 SV=1 - [A6NNI4_HUMAN]	1	1	1
P20160	Azurocidin OS=Homo sapiens GN=AZU1 PE=1 SV=3 - [CAP7_HUMAN]	1	1	1
G1TVL1	Elongation factor 1-gamma OS=Oryctolagus cuniculus GN=EEF1G PE=4 SV=1 - [G1TVL1_RABIT]	2	2	2
P98160	Basement membrane-specific heparan sulfate proteoglycan core protein OS=Homo sapiens GN=HSPG2 PE=	1	1	1
P01704	Immunoglobulin lambda variable 2-14 OS=Homo sapiens GN=IGLV2-14 PE=1 SV=2 - [LV214_HUMAN]	1	1	3
P50454	Serpin H1 OS=Homo sapiens GN=SERPINH1 PE=1 SV=2 - [SERPH_HUMAN]	2	2	2
C9JD14	Guanine nucleotide-binding protein subunit beta-4 (Fragment) OS=Homo sapiens GN=GNB4 PE=1 SV=1 - [1	1	1
G1T3Y8	Uncharacterized protein OS=Oryctolagus cuniculus GN=HSPD1 PE=3 SV=2 - [G1T3Y8_RABIT]	1	1	1
B4DYT9	Sarcalumenin OS=Homo sapiens GN=SRL PE=1 SV=1 - [B4DYT9_HUMAN]	1	1	1
P60903	Protein S100-A10 OS=Homo sapiens GN=S100A10 PE=1 SV=2 - [S10AA_HUMAN]	1	1	1

Q6YHK3	CD109 antigen OS=Homo sapiens GN=CD109 PE=1 SV=2 - [CD109_HUMAN]	2	2	2
P10301	Ras-related protein R-Ras OS=Homo sapiens GN=RRAS PE=1 SV=1 - [RRAS_HUMAN]	2	2	2
H0Y4U1	Tensin-1 OS=Homo sapiens GN=TNS1 PE=1 SV=1 - [H0Y4U1_HUMAN]	1	1	1
Q5VVQ7	C4b-binding protein beta chain (Fragment) OS=Homo sapiens GN=C4BPB PE=1 SV=1 - [Q5VVQ7_HUMAN]	1	1	1
P00740	Coagulation factor IX OS=Homo sapiens GN=F9 PE=1 SV=2 - [FA9_HUMAN]	1	1	1
P30357	Arachidonate 5-lipoxygenase-activating protein (Fragment) OS=Oryctolagus cuniculus GN=ALOX5AP PE=2	1	1	1
H7C123	60S ribosomal protein L10 (Fragment) OS=Homo sapiens GN=RPL10 PE=1 SV=1 - [H7C123_HUMAN]	1	1	1
A0A0C4DH35	Protein IGHV3-35 (Fragment) OS=Homo sapiens GN=IGHV3-35 PE=1 SV=1 - [A0A0C4DH35_HUMAN]	1	1	1
GIU6T1	Carboxylic ester hydrolase OS=Oryctolagus cuniculus GN=LOC100355283 PE=3 SV=1 - [GIU6T1_RABIT]	1	1	1
P03973	Antileukoproteinase OS=Homo sapiens GN=SLPI PE=1 SV=2 - [SLPI_HUMAN]	1	1	1
D6RD47	40S ribosomal protein S23 OS=Homo sapiens GN=RPS23 PE=1 SV=1 - [D6RD47_HUMAN]	1	1	1
A6PVX3	26S proteasome non-ATPase regulatory subunit 4 (Fragment) OS=Homo sapiens GN=PSMD4 PE=1 SV=1 -	1	1	1
GI1TY53	Ribosomal protein OS=Oryctolagus cuniculus PE=3 SV=1 - [GI1TY53_RABIT]	1	1	1
A0A0A0MS15	Protein IGHV3-49 (Fragment) OS=Homo sapiens GN=IGHV3-49 PE=1 SV=1 - [A0A0A0MS15_HUMAN]	1	2	2
P12314	High affinity immunoglobulin gamma Fc receptor 1 OS=Homo sapiens GN=FCGR1A PE=1 SV=2 - [FCGR1_F	2	2	2
A0A075B6J9	HCG1782423 (Fragment) OS=Homo sapiens GN=IGLV2-18 PE=1 SV=1 - [A0A075B6J9_HUMAN]	3	3	3
X6RLJ0	Complement C1q subcomponent subunit A (Fragment) OS=Homo sapiens GN=C1QA PE=1 SV=7 - [X6RLJ0	1	1	1
MQQZ2	40S ribosomal protein S5 OS=Homo sapiens GN=RPS5 PE=1 SV=1 - [MQQZ2_HUMAN]	1	1	1
GI1TME7	Uncharacterized protein OS=Oryctolagus cuniculus GN=CSTB PE=4 SV=1 - [GI1TME7_RABIT]	1	1	1
E9PQR3	Ferritin OS=Homo sapiens GN=FTTH1 PE=1 SV=1 - [E9PQR3_HUMAN]	1	1	1
P01008	Antithrombin-III OS=Homo sapiens GN=SERPINC1 PE=1 SV=1 - [ANT3_HUMAN]	2	2	2
A0A087WSY5	Carboxypeptidase B2 OS=Homo sapiens GN=CPB2 PE=1 SV=1 - [A0A087WSY5_HUMAN]	1	1	1
K7EP04	Heat shock protein beta-6 OS=Homo sapiens GN=HSPB6 PE=1 SV=1 - [K7EP04_HUMAN]	1	1	1
B8Z251	Malate dehydrogenase, cytoplasmic OS=Homo sapiens GN=MDH1 PE=1 SV=1 - [B8Z251_HUMAN]	1	1	1
P62318	Small nuclear ribonucleoprotein Sm D3 OS=Homo sapiens GN=SNRPD3 PE=1 SV=1 - [SMD3_HUMAN]	1	1	1
C9JJV6	Myeloid-associated differentiation marker (Fragment) OS=Homo sapiens GN=MYADM PE=1 SV=8 - [C9JJV6	1	1	1
C9JCA5	Troponin T, fast skeletal muscle (Fragment) OS=Homo sapiens GN=TNNT3 PE=1 SV=1 - [C9JCA5_HUMAN]	2	2	2
E9PKC6	CD44 antigen OS=Homo sapiens GN=CD44 PE=1 SV=1 - [E9PKC6_HUMAN]	1	1	1
P27449	V-type proton ATPase 16 kDa proteolipid subunit OS=Homo sapiens GN=ATP6V0C PE=1 SV=1 - [VATL_HU	1	1	2
GI1T3W0	Uncharacterized protein OS=Oryctolagus cuniculus GN=LDB3 PE=4 SV=2 - [GI1T3W0_RABIT]	1	1	1
GI1SND0	Sulphydryl oxidase OS=Oryctolagus cuniculus GN=QSOX1 PE=4 SV=2 - [GI1SND0_RABIT]	1	1	1
Q5JX18	Four and a half LIM domains protein 1 (Fragment) OS=Homo sapiens GN=FHL1 PE=1 SV=1 - [Q5JX18_HUI	3	3	3
Q9UBR2	Cathepsin Z OS=Homo sapiens GN=CTSZ PE=1 SV=1 - [CATZ_HUMAN]	1	1	1
H0YLP6	60S ribosomal protein L28 OS=Homo sapiens GN=RPL28 PE=1 SV=1 - [H0YLP6_HUMAN]	2	2	2
P49913	Cathelicidin antimicrobial peptide OS=Homo sapiens GN=CAMP PE=1 SV=1 - [CAMP_HUMAN]	1	1	1
GI1U7S4	Phosphoglycerate mutase OS=Oryctolagus cuniculus GN=PGAM2 PE=3 SV=1 - [GI1U7S4_RABIT]	2	2	2
P35579	Myosin-9 OS=Homo sapiens GN=MYH9 PE=1 SV=4 - [MYH9_HUMAN]	1	4	4
GI1TC33	Uncharacterized protein OS=Oryctolagus cuniculus GN=MYH10 PE=3 SV=2 - [GI1TC33_RABIT]	1	4	4
P61106	Ras-related protein Rab-14 OS=Homo sapiens GN=RAB14 PE=1 SV=4 - [RAB14_HUMAN]	1	1	1
P02765	Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1 - [FETUA_HUMAN]	1	1	1
GI1S254	Uncharacterized protein OS=Oryctolagus cuniculus GN=FRZB PE=4 SV=2 - [GI1S254_RABIT]	1	1	1
P40261	Nicotinamide N-methyltransferase OS=Homo sapiens GN=NNMT PE=1 SV=1 - [NNMT_HUMAN]	3	3	3
P61026	Ras-related protein Rab-10 OS=Homo sapiens GN=RAB10 PE=1 SV=1 - [RAB10_HUMAN]	1	1	1

R4GN98	Protein S100 (Fragment) OS=Homo sapiens GN=S100A6 PE=1 SV=1 - [R4GN98_HUMAN]	1	1	1
P24752	Acetyl-CoA acetyltransferase, mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1 - [THIL_HUMAN]	1	1	1
P21195	Protein disulfide-isomerase OS=Oryctolagus cuniculus GN=P4HB PE=2 SV=1 - [PDIA1_RABIT]	4	4	4
Q5T750	Skin-specific protein 32 OS=Homo sapiens GN=XP32 PE=1 SV=1 - [XP32_HUMAN]	1	1	1
G1T277	Actin-related protein 2/3 complex subunit 3 OS=Oryctolagus cuniculus GN=ARPC3 PE=3 SV=1 - [G1T277_HUMAN]	3	3	3
P11279	Lysosome-associated membrane glycoprotein 1 OS=Homo sapiens GN=LAMP1 PE=1 SV=3 - [LAMP1_HUMAN]	1	1	1
G1T235	Proteasome subunit beta type OS=Oryctolagus cuniculus GN=PSMB6 PE=3 SV=1 - [LAMP1_HUMAN]	1	1	1
P31347	Angiogenin OS=Oryctolagus cuniculus GN=ANG PE=1 SV=1 - [ANG1_RABBIT]	1	1	1
F2Z393	Transaldolase OS=Homo sapiens GN=TALDO1 PE=1 SV=1 - [F2Z393_HUMAN]	1	1	1
Q8WUIM4	Programmed cell death 6-interacting protein OS=Homo sapiens GN=PDC6I1 PE=1 SV=1 - [PDC6I_HUMAN]	3	3	3
Q16270	Insulin-like growth factor-binding protein 7 OS=Homo sapiens GN=IGFBP7 PE=1 SV=1 - [IBP7_HUMAN]	2	2	2
D6RA16	Filamin-binding LIM protein 1 (Fragment) OS=Homo sapiens GN=FBLIM1 PE=1 SV=1 - [D6RA16_HUMAN]	2	2	2
G1T7R2	Uncharacterized protein OS=Oryctolagus cuniculus GN=YWHAE PE=3 SV=1 - [G1T7R2_RABIT]	2	2	4
G1TZP0	Uncharacterized protein OS=Oryctolagus cuniculus GN=YWHAG PE=3 SV=1 - [G1TZP0_RABIT]	2	2	5
K7EMN2	6-phosphogluconate dehydrogenase, decarboxylating (Fragment) OS=Homo sapiens GN=PGD PE=1 SV=1	1	1	1
H0Y4R1	Inosine 5'-monophosphate dehydrogenase 2 (Fragment) OS=Homo sapiens GN=IMPDH2 PE=1 SV=1 - [HC]	1	1	1
V9GYE3	Apolipoprotein A-II OS=Homo sapiens GN=APOA2 PE=1 SV=1 - [V9GYE3_HUMAN]	1	1	1
Q04941	Proteolipid protein 2 OS=Homo sapiens GN=PLP2 PE=1 SV=1 - [PLP2_HUMAN]	1	1	1
Q9HB40	Retinoid-inducible serine carboxypeptidase OS=Homo sapiens GN=SCEP1 PE=1 SV=1 - [RISC_HUMAN]	1	1	1
E7EX53	Ribosomal protein L15 (Fragment) OS=Homo sapiens GN=RPL15 PE=1 SV=1 - [E7EX53_HUMAN]	1	1	1
A0A0B4J1V2	Protein IGHV2-26 (Fragment) OS=Homo sapiens GN=IGHV2-26 PE=1 SV=1 - [A0A0B4J1V2_HUMAN]	1	1	1
Q9Y3Z3	Deoxynucleoside triphosphate triphosphohydrolase SAMHD1 OS=Homo sapiens GN=SAMHD1 PE=1 SV=2	1	1	1
P08571	Monocyte differentiation antigen CD14 OS=Homo sapiens GN=CD14 PE=1 SV=2 - [CD14_HUMAN]	2	2	2
A0A0A0MT69	Protein IGK34 (Fragment) OS=Homo sapiens GN=IGK4 PE=1 SV=1 - [A0A0A0MT69_HUMAN]	1	1	1
D6RGG3	Collagen alpha-1(XI) chain OS=Homo sapiens GN=COL12A1 PE=1 SV=1 - [D6RGG3_HUMAN]	3	3	3
G1T7A8	Uncharacterized protein OS=Oryctolagus cuniculus GN=CPQ PE=4 SV=1 - [G1T7A8_RABIT]	1	1	1
P04196	Histidine-rich glycoprotein OS=Homo sapiens GN=HRG PE=1 SV=1 - [HRG_HUMAN]	2	2	2
H0Y984	ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 2 (Fragment) OS=Homo sapiens GN=BST1 PE=1 SV=1 - [I]	1	1	1
P10253	Lysosomal alpha-glucosidase OS=Homo sapiens GN=GAA PE=1 SV=4 - [LYAG_HUMAN]	1	1	1
F5H316	Fermitin family homolog 3 (Fragment) OS=Homo sapiens GN=FERMT3 PE=1 SV=1 - [F5H316_HUMAN]	1	1	1
O14791	Apolipoprotein L1 OS=Homo sapiens GN=APOL1 PE=1 SV=5 - [APOL1_HUMAN]	1	1	1
P04040	Catalase OS=Homo sapiens GN=CAT PE=1 SV=3 - [CATA_HUMAN]	3	3	3
P01705	Immunoglobulin lambda variable 2-23 OS=Homo sapiens GN=IGLV2-23 PE=1 SV=2 - [LV23_HUMAN]	1	1	1
K7EQQ9	60S ribosomal protein L27 OS=Homo sapiens GN=RPL27 PE=1 SV=1 - [K7EQQ9_HUMAN]	2	2	2
M0QX76	40S ribosomal protein S16 (Fragment) OS=Homo sapiens GN=RPS16 PE=1 SV=1 - [M0QX76_HUMAN]	1	1	1
A0A0D9SG66	Isocitrate dehydrogenase [NAD] subunit beta, mitochondrial (Fragment) OS=Homo sapiens GN=IDH3B PE=1 SV=1 - [ASPN_HUMAN]	1	1	1
Q9BXN1	Asporin OS=Homo sapiens GN=ASPN PE=1 SV=2 - [ASPN_HUMAN]	1	1	1
Q5QNZ2	ATP synthase F0 complex subunit B1, mitochondrial OS=Homo sapiens GN=ATP5F1 PE=1 SV=1 - [Q5QNZ2]	1	1	1
G1T9M7	Alpha-1,4 glucan phosphorylase OS=Oryctolagus cuniculus GN=PYGL PE=3 SV=2 - [G1T9M7_RABIT]	1	3	3
G1SGF8	Uncharacterized protein OS=Oryctolagus cuniculus GN=C6 PE=4 SV=2 - [G1SGF8_RABIT]	1	1	1
E7ESS4	Intercellular adhesion molecule 1 OS=Homo sapiens GN=ICAM1 PE=1 SV=1 - [E7ESS4_HUMAN]	1	1	1
P62851	40S ribosomal protein S25 OS=Homo sapiens GN=RPS25 PE=1 SV=1 - [RS25_HUMAN]	1	1	1
P01833	Polymeric immunoglobulin receptor OS=Homo sapiens GN=PIGR PE=1 SV=4 - [PIGR_HUMAN]	1	1	1

Q5D862	Filaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 - [FILA2_HUMAN]	1	1	1
G1U7K4	Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=1 - [G1U7K4_RABIT]	1	1	1
G1T013	Malic enzyme OS=Oryctolagus cuniculus GN=ME1 PE=3 SV=2 - [G1T013_RABIT]	1	1	1
P29622	Kallistatin OS=Homo sapiens GN=SERPINA4 PE=1 SV=3 - [KAIN_HUMAN]	1	1	1
J3KRG2	Gasdermin-A (Fragment) OS=Homo sapiens GN=GSDMA PE=1 SV=8 - [J3KRG2_HUMAN]	1	1	1
P80108	Phosphatidylinositol-glycan-specific phospholipase D OS=Homo sapiens GN=GPLD1 PE=1 SV=3 - [PHLD_HUMAN]	1	1	1
A0A0U1RQV3	EGF-containing fibulin-like extracellular matrix protein 1 (Fragment) OS=Homo sapiens GN=EFEMP1 PE=1 SV=1 - [H0Y8Y3_HUMAN]	3	3	3
H0Y8Y3	PDZ and LIM domain protein 5 (Fragment) OS=Homo sapiens GN=PDLIM5 PE=1 SV=1 - [H0Y8Y3_HUMAN]	1	1	1
P13473	Lysosome-associated membrane glycoprotein 2 OS=Homo sapiens GN=LAMP2 PE=1 SV=2 - [LAMP2_HUMAN]	1	1	1
P31944	Caspase-14 OS=Homo sapiens GN=CASP14 PE=1 SV=2 - [CASPE_HUMAN]	2	2	2
A0A087WX18	HLA class II histocompatibility antigen, DRB1-16 beta chain OS=Homo sapiens GN=HLA-DRB1 PE=1 SV=1	2	2	2
P30041	Peroxioredoxin-6 OS=Homo sapiens GN=PRDX6 PE=1 SV=3 - [PRDX6_HUMAN]	1	1	1
P05164	Myeloperoxidase OS=Homo sapiens GN=MPO PE=1 SV=1 - [PERM_HUMAN]	3	3	3
G1SRF7	Uncharacterized protein OS=Oryctolagus cuniculus GN=HSPA9 PE=3 SV=1 - [G1SRF7_RABIT]	2	2	2
G1TNN3	Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=1 - [G1TNN3_RABIT]	1	1	1
P08603	Complement factor H OS=Homo sapiens GN=CFH PE=1 SV=4 - [CFAH_HUMAN]	1	1	1
P35858	Insulin-like growth factor-binding protein complex acid labile subunit OS=Homo sapiens GN=IGFALS PE=1	2	2	2
C9JUG7	F-actin-capping protein subunit alpha-2 OS=Homo sapiens GN=CAPZA2 PE=1 SV=1 - [C9JUG7_HUMAN]	1	1	1
Q07960	Rho GTPase-activating protein 1 OS=Homo sapiens GN=ARHGAP1 PE=1 SV=1 - [RHG01_HUMAN]	1	1	1
G1SHY5	Uncharacterized protein OS=Oryctolagus cuniculus GN=HAPLN1 PE=4 SV=1 - [G1SHY5_RABIT]	1	1	1
P30044	Peroxioredoxin-5, mitochondrial OS=Homo sapiens GN=PRDX5 PE=1 SV=4 - [PRDX5_HUMAN]	1	1	1
G1SIT6	Actin-related protein 2/3 complex subunit 5 OS=Oryctolagus cuniculus GN=ARPC5L PE=3 SV=1 - [G1SIT6_HUMAN]	1	1	1
K7EMV3	Histone H3 OS=Homo sapiens GN=H3F3B PE=1 SV=1 - [K7EMV3_HUMAN]	1	1	1
P50238	Cysteine-rich protein 1 OS=Homo sapiens GN=CRIP1 PE=1 SV=3 - [CRIP1_HUMAN]	2	2	2
E7EMC6	Annexin OS=Homo sapiens GN=ANXA6 PE=1 SV=1 - [E7EMC6_HUMAN]	1	1	1
G1TUC8	Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=2 - [G1TUC8_RABIT]	1	3	3
A0A087WTS8	Heat shock 70 kDa protein 4 OS=Homo sapiens GN=HSPA4 PE=1 SV=1 - [A0A087WTS8_HUMAN]	1	1	1
Q5SX91	Rab GDP dissociation inhibitor beta (Fragment) OS=Homo sapiens GN=GD12 PE=1 SV=1 - [Q5SX91_HUMAN]	1	1	1
F5GY37	Prohibitin-2 OS=Homo sapiens GN=PHB2 PE=1 SV=1 - [F5GY37_HUMAN]	2	2	2
P15814	Immunoglobulin lambda-like polypeptide 1 OS=Homo sapiens GN=IGLL1 PE=1 SV=1 - [IGLL1_HUMAN]	1	1	1
Q15149	Plectin OS=Homo sapiens GN=PLEC PE=1 SV=3 - [PLEC_HUMAN]	1	1	1
G1SGH0	Uncharacterized protein OS=Oryctolagus cuniculus GN=LOC100339859 PE=3 SV=1 - [G1SGH0_RABIT]	1	1	4

Accession	Description	# Unique Peptides	# Peptides	# PSMs
P01857	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1	11	22	142
A0A075B6N8	Ig gamma-3 chain C region (Fragment) OS=Homo sapiens GN=IG	6	18	125
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1	34	41	120
P04114	Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=2	106	106	135
P01859	Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1	9	17	114
P60709	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1	11	25	126
P01834	Ig kappa chain C region OS=Homo sapiens GN=IGKC PE=1 SV=1	8	8	95
P62736	Actin, aortic smooth muscle OS=Homo sapiens GN=ACTA2 PE=1	11	25	121
P21333	Filamin-A OS=Homo sapiens GN=FLNA PE=1 SV=4	80	80	98
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=	29	35	93
A0A075B6K9	Ig lambda-2 chain C regions (Fragment) OS=Homo sapiens GN=IC	2	6	67
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KI	27	35	62
B9A064	Immunoglobulin lambda-like polypeptide 5 OS=Homo sapiens GN=	3	6	56
P01861	Ig gamma-4 chain C region OS=Homo sapiens GN=IGHG4 PE=1	5	12	54
P08779	Keratin, type I cytoskeletal 16 OS=Homo sapiens GN=KRT16 PE=	12	27	61
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1	28	29	64
P02533	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=	6	24	55
P08670	Vimentin OS=Homo sapiens GN=VIM PE=1 SV=4	34	36	57
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1	20	33	46
P02538	Keratin, type II cytoskeletal 6A OS=Homo sapiens GN=KRT6A PE=	3	34	45
P04259	Keratin, type II cytoskeletal 6B OS=Homo sapiens GN=KRT6B PE=	2	33	45
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	28	28	45
A0A075B6L1	Ig lambda-7 chain C region (Fragment) OS=Homo sapiens GN=IG	1	4	34
E7ENL6	Collagen alpha-3(VI) chain OS=Homo sapiens GN=COL6A3 PE=1	40	40	43
A0A087WYJ9	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=2	7	16	36
Q562R1	Beta-actin-like protein 2 OS=Homo sapiens GN=ACTBL2 PE=1 SV=	1	8	62
P10909	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1	18	18	35
P02649	Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1	20	20	34
P51888	Prolargin OS=Homo sapiens GN=PRELP PE=1 SV=1	16	16	31
P21810	Biglycan OS=Homo sapiens GN=BGN PE=1 SV=2	15	15	24
P02751	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=4	34	34	38
P04004	Vitronectin OS=Homo sapiens GN=VTN PE=1 SV=1	11	11	26
P02647	Apolipoprotein A-I OS=Homo sapiens GN=APOA1 PE=1 SV=1	18	18	27
P68871	Hemoglobin subunit beta OS=Homo sapiens GN=HBB PE=1 SV=2	11	11	23
P04406	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens G	17	17	24
P04220	Ig mu heavy chain disease protein OS=Homo sapiens PE=1 SV=1	1	10	21
P01024	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	26	26	27
P01876	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=	4	11	20

E7EQB2	Lactotransferrin (Fragment) OS=Homo sapiens GN=LTF PE=1 SV=	22	22	24
P14618	Pyruvate kinase PKM OS=Homo sapiens GN=PKM PE=1 SV=4 - [K	18	18	24
Q04695	Keratin, type I cytoskeletal 17 OS=Homo sapiens GN=KRT17 PE=	5	15	21
P02675	Fibrinogen beta chain OS=Homo sapiens GN=FGB PE=1 SV=2 - [I	18	18	24
P01009	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	18	18	22
Q08431	Lactadherin OS=Homo sapiens GN=MFG8 PE=1 SV=2 - [MFGM_	2	14	20
A0A087X0S5	Collagen alpha-1(VI) chain OS=Homo sapiens GN=COL6A1 PE=1	14	14	18
P13611	Versican core protein OS=Homo sapiens GN=VCAN PE=1 SV=3 -	6	16	19
X6R3G6	Lactadherin OS=Homo sapiens GN=MFG8 PE=1 SV=1 - [X6R3G6	1	13	18
Q71U36	Tubulin alpha-1A chain OS=Homo sapiens GN=TUBA1A PE=1 SV=	1	12	19
P68363	Tubulin alpha-1B chain OS=Homo sapiens GN=TUBA1B PE=1 SV=	2	13	20
G1U6W1	Elongation factor 1-alpha OS=Oryctolagus cuniculus PE=3 SV=1 -	11	11	17
P02743	Serum amyloid P-component OS=Homo sapiens GN=APCS PE=1 S	10	10	15
P01023	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=3 -	20	20	22
P07437	Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV=2 - [TF	4	13	17
A0A0G2JMB2	Ig alpha-2 chain C region (Fragment) OS=Homo sapiens GN=IGH/	1	8	14
C9JEU5	Fibrinogen gamma chain OS=Homo sapiens GN=FGG PE=1 SV=1	12	12	17
D6RGZ6	Versican core protein (Fragment) OS=Homo sapiens GN=VCAN PE	1	11	14
A0A0C4DGL8	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1 - [A0A0C4DG	5	12	15
Q05707	Collagen alpha-1(XIV) chain OS=Homo sapiens GN=COL14A1 PE=	15	15	16
P01031	Complement C5 OS=Homo sapiens GN=C5 PE=1 SV=4 - [CO5_HI	14	14	15
P01619	Immunoglobulin kappa variable 3-20 OS=Homo sapiens GN=IGKV	2	5	9
P07355	Annexin A2 OS=Homo sapiens GN=ANXA2 PE=1 SV=2 - [ANXA2_	10	10	10
A0A1B0GW44	Cathepsin D OS=Homo sapiens GN=CTSD PE=1 SV=1 - [A0A1B0C	12	12	16
E9PKE3	Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8	9	11	13
G1T6D3	Uncharacterized protein OS=Oryctolagus cuniculus GN=KRT72 PE=	1	6	10
P07360	Complement component C8 gamma chain OS=Homo sapiens GN=	6	6	7
Q13885	Tubulin beta-2A chain OS=Homo sapiens GN=TUBB2A PE=1 SV=	1	10	14
Q7Z794	Keratin, type II cytoskeletal 1b OS=Homo sapiens GN=KRT77 PE=	4	7	11
P04792	Heat shock protein beta-1 OS=Homo sapiens GN=HSPB1 PE=1 SV	9	9	11
Q02413	Desmoglein-1 OS=Homo sapiens GN=DSG1 PE=1 SV=2 - [DSG1_	10	10	10
P06312	Immunoglobulin kappa variable 4-1 OS=Homo sapiens GN=IGKV4	3	3	9
P51884	Lumican OS=Homo sapiens GN=LUM PE=1 SV=2 - [LUM_HUMAN	9	9	12
P10301	Ras-related protein R-Ras OS=Homo sapiens GN=RRAS PE=1 SV=	8	8	9
P23284	Peptidyl-prolyl cis-trans isomerase B OS=Homo sapiens GN=PP1B	10	10	13
B1ALD9	Periostin OS=Homo sapiens GN=POSTN PE=1 SV=1 - [B1ALD9_H	12	12	13
P00739	Haptoglobin-related protein OS=Homo sapiens GN=HPR PE=2 SV=	4	11	11
P69905	Hemoglobin subunit alpha OS=Homo sapiens GN=HBA1 PE=1 SV=	5	5	9
A0A0A0MSV6	Complement C1q subcomponent subunit B (Fragment) OS=Homo	6	6	9

Q08380	Galectin-3-binding protein OS=Homo sapiens GN=LGALS3BP PE=1	12	12	13
A0A087X0Q4	Protein IGV2-40 OS=Homo sapiens GN=IGKV2-40 PE=1 SV=1 -	1	3	6
P12110	Collagen alpha-2(VI) chain OS=Homo sapiens GN=COL6A2 PE=1	9	9	9
P00338	L-lactate dehydrogenase A chain OS=Homo sapiens GN=LDHA PE	9	10	10
A0A140TA32	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1 - [A0A	1	12	14
P04003	C4b-binding protein alpha chain OS=Homo sapiens GN=C4BPA PE	9	9	10
P00325	Alcohol dehydrogenase 1B OS=Homo sapiens GN=ADH1B PE=1 S	9	9	10
A0A0J9YY99	Uncharacterized protein (Fragment) OS=Homo sapiens PE=1 SV=	4	5	8
A0A087WW89	Protein IGHV3-72 OS=Homo sapiens GN=IGHV3-72 PE=1 SV=1 -	2	3	7
A0A0C4DH25	Immunoglobulin kappa variable 3-20 OS=Homo sapiens GN=IGKV	1	4	7
P06576	ATP synthase subunit beta, mitochondrial OS=Homo sapiens GN=	9	9	10
P15924	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3 - [DESP_HU	12	12	12
A0A0G2JL54	Complement C4-B OS=Homo sapiens GN=C4B_2 PE=1 SV=1 - [AI	2	13	15
A0A087X0P6	Protein IGV2D-29 OS=Homo sapiens GN=IGKV2D-29 PE=1 SV=	1	4	6
P02747	Complement C1q subcomponent subunit C OS=Homo sapiens GN=	5	5	5
G1SHZ4	Uncharacterized protein OS=Oryctolagus cuniculus GN=KRT7 PE=	1	5	8
P63244	Receptor of activated protein C kinase 1 OS=Homo sapiens GN=R.	8	8	9
P02671	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2 -	9	9	11
P01825	Immunoglobulin heavy variable 4-59 OS=Homo sapiens GN=IGHV	2	3	6
P02749	Beta-2-glycoprotein 1 OS=Homo sapiens GN=APOH PE=1 SV=3 -	6	6	7
A0A075B6K5	HCG2043239 (Fragment) OS=Homo sapiens GN=IGLV3-9 PE=1 S	3	3	8
P05090	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1 - [APC	6	6	9
P25705	ATP synthase subunit alpha, mitochondrial OS=Homo sapiens GN=	7	7	7
P08519	Apolipoprotein(a) OS=Homo sapiens GN=LPA PE=1 SV=1 - [APOA	6	6	7
P06727	Apolipoprotein A-IV OS=Homo sapiens GN=APOA4 PE=1 SV=3 - [6	6	7
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4 OS=Homo sapiens GN	6	6	7
P35749	Myosin-11 OS=Homo sapiens GN=MYH11 PE=1 SV=3 - [MYH11_	6	9	9
P07093	Glia-derived nexin OS=Homo sapiens GN=SERPINE2 PE=1 SV=1 -	5	5	6
P35542	Serum amyloid A-4 protein OS=Homo sapiens GN=SA4A PE=1 SV	4	4	6
Q9Y490	Talin-1 OS=Homo sapiens GN=TLN1 PE=1 SV=3 - [TLN1_HUMAN	8	8	8
P01624	Immunoglobulin kappa variable 3-15 OS=Homo sapiens GN=IGKV	1	2	5
P0DMV8	Heat shock 70 kDa protein 1A OS=Homo sapiens GN=HSPA1A PE=	8	9	10
Q92743	Serine protease HTRA1 OS=Homo sapiens GN=HTRA1 PE=1 SV=	5	5	5
P01718	Immunoglobulin lambda variable 3-27 OS=Homo sapiens GN=IGL	4	4	5
P62820	Ras-related protein Rab-1A OS=Homo sapiens GN=RAB1A PE=1 S	7	9	9
Q01995	Transgelin OS=Homo sapiens GN=TAGLN PE=1 SV=4 - [TAGL_HU	6	6	6
P61626	Lysozyme C OS=Homo sapiens GN=LYZ PE=1 SV=1 - [LYSC_HUM	5	5	5
A0A075B7B8	Protein IGHV3OR16-12 (Fragment) OS=Homo sapiens GN=IGHV3	1	2	5
G1T2J0	Uncharacterized protein OS=Oryctolagus cuniculus GN=JUP PE=4	6	6	6

Q5T985	Inter-alpha-trypsin inhibitor heavy chain H2 OS=Homo sapiens GN	6	6	8
P04432	Immunoglobulin kappa variable 1D-39 OS=Homo sapiens GN=IGK	2	2	4
P19827	Inter-alpha-trypsin inhibitor heavy chain H1 OS=Homo sapiens GN	6	6	6
P01782	Immunoglobulin heavy variable 3-9 OS=Homo sapiens GN=IGHV3	2	4	8
P35858	Insulin-like growth factor-binding protein complex acid labile subur	6	6	6
Q8IUX7	Adipocyte enhancer-binding protein 1 OS=Homo sapiens GN=AEB1	5	5	5
P01701	Immunoglobulin lambda variable 1-51 OS=Homo sapiens GN=IGL1	5	5	5
P00558	Phosphoglycerate kinase 1 OS=Homo sapiens GN=PGK1 PE=1 SV=	6	6	6
AOA075B6S5	Protein IGKV1-27 (Fragment) OS=Homo sapiens GN=IGKV1-27 PE	1	2	3
AOA075B6R2	Protein IGHV4-4 (Fragment) OS=Homo sapiens GN=IGHV4-4 PE=	2	3	6
AOA0C4DH67	Protein IGKV1-8 (Fragment) OS=Homo sapiens GN=IGKV1-8 PE=	1	2	3
P07195	L-lactate dehydrogenase B chain OS=Homo sapiens GN=LDHB PE=	4	5	5
P01042	Kinogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2 - [KNG1_H	7	7	7
P02760	Protein AMBP OS=Homo sapiens GN=AMBP PE=1 SV=1 - [AMBP_	5	5	6
E9PHK0	Tetranectin OS=Homo sapiens GN=CLEC3B PE=1 SV=1 - [E9PHK	4	4	4
AOA0G2JQR6	Uncharacterized protein (Fragment) OS=Homo sapiens PE=1 SV=	3	3	5
P07737	Profilin-1 OS=Homo sapiens GN=PFN1 PE=1 SV=2 - [PROF1_HUJ	5	5	5
P06733	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [ENOA_	5	5	6
P48745	Protein NOV homolog OS=Homo sapiens GN=NOV PE=1 SV=1 - [6	6	6
Q8N1N4	Keratin, type II cytoskeletal 78 OS=Homo sapiens GN=KRT78 PE=	3	4	5
AOA0A0MRZ8	Protein IGKV3D-11 (Fragment) OS=Homo sapiens GN=IGKV3D-11	2	2	6
P60174	Triosephosphate isomerase OS=Homo sapiens GN=TPI1 PE=1 SV	5	5	5
P08133	Annexin A6 OS=Homo sapiens GN=ANXA6 PE=1 SV=3 - [ANXA6_	5	5	5
P01599	Immunoglobulin kappa variable 1-17 OS=Homo sapiens GN=IGKV	3	3	3
P01700	Immunoglobulin lambda variable 1-47 OS=Homo sapiens GN=IGL	1	2	4
AOA0A0MSI0	Peroxiredoxin-1 (Fragment) OS=Homo sapiens GN=PRDX1 PE=1	2	5	5
AOA075B6H7	Protein IGKV3-7 (Fragment) OS=Homo sapiens GN=IGKV3-7 PE=	1	2	5
AOA0D9SFZ9	Four and a half LIM domains protein 1 (Fragment) OS=Homo sapi	3	3	4
Q14195	Dihydropyrimidinase-related protein 3 OS=Homo sapiens GN=DPY	4	5	5
P07996	Thrombospondin-1 OS=Homo sapiens GN=THBS1 PE=1 SV=2 - [5	5	5
P01780	Immunoglobulin heavy variable 3-7 OS=Homo sapiens GN=IGHV3	2	4	7
Q96PD5	N-acetylmuramoyl-L-alanine amidase OS=Homo sapiens GN=PGLY	4	4	4
P01011	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1	4	4	4
P27169	Serum paraoxonase/arylesterase 1 OS=Homo sapiens GN=PON1 f	5	5	5
Q06828	Fibromodulin OS=Homo sapiens GN=FMOD PE=1 SV=2 - [FMOD_	4	4	5
P60981	Destrin OS=Homo sapiens GN=DSTN PE=1 SV=3 - [DEST_HUMAI	5	5	5
G1TAX7	Uncharacterized protein OS=Oryctolagus cuniculus GN=RAC2 PE=	4	5	5
C9J592	Ras-related protein Rab-7a (Fragment) OS=Homo sapiens GN=RA	4	4	4
C9J1Z8	ADP-ribosylation factor 5 (Fragment) OS=Homo sapiens GN=ARF5	3	3	4

P29508	Serpin B3 OS=Homo sapiens GN=SERPINB3 PE=1 SV=2 - [SPB3_	5	5	5	5
F8VQR7	Cysteine and glycine-rich protein 2 OS=Homo sapiens GN=CSR2	4	4	4	4
F8VPF3	Myosin light polypeptide 6 (Fragment) OS=Homo sapiens GN=MYI	5	5	5	6
Q15365	Poly(rC)-binding protein 1 OS=Homo sapiens GN=PCBP1 PE=1 SV	4	4	5	5
H9KV48	Plasma protease C1 inhibitor OS=Homo sapiens GN=SERPING1 PE	4	4	4	5
P25311	Zinc-alpha-2-glycoprotein OS=Homo sapiens GN=AZGP1 PE=1 SV	6	6	6	6
F5H018	GTP-binding nuclear protein Ran (Fragment) OS=Homo sapiens G	4	4	4	5
P01602	Immunoglobulin kappa variable 1-5 OS=Homo sapiens GN=IGKV1	2	2	2	4
A0A0C4DH38	Protein IGHV5-51 (Fragment) OS=Homo sapiens GN=IGHV5-51 P	2	2	2	3
A0A0A0MS51	Gelsolin OS=Homo sapiens GN=GSN PE=1 SV=1 - [A0A0A0MS51_	6	6	6	6
P04899	Guanine nucleotide-binding protein G(i) subunit alpha-2 OS=Homc	2	3	3	3
P01714	Immunoglobulin lambda variable 3-19 OS=Homo sapiens GN=IGL	2	2	2	4
A0A087WV45	Transferrin OS=Homo sapiens GN=TTR PE=1 SV=1 - [A0A087V	4	4	4	4
A0A087X0A8	Collagen alpha-1(XII) chain OS=Homo sapiens GN=COL12A1 PE=	6	6	6	6
P98160	Basement membrane-specific heparan sulfate proteoglycan core pi	5	5	5	5
Q6PID9	Aggrecan OS=Homo sapiens GN=ACAN PE=1 SV=1 - [Q6PID9_H	4	4	4	4
Q86YZ3	Hormerin OS=Homo sapiens GN=HRNR PE=1 SV=2 - [HORN_HUM	2	2	2	2
P61026	Ras-related protein Rab-10 OS=Homo sapiens GN=RAB10 PE=1 S	1	1	3	3
P23083	Immunoglobulin heavy variable 1-2 OS=Homo sapiens GN=IGHV1	1	1	2	2
P00747	Plasminogen OS=Homo sapiens GN=PLG PE=1 SV=2 - [PLMN_HU	7	7	7	7
G1TKY9	Uncharacterized protein OS=Oryctolagus cuniculus GN=CTHRC1 P	3	3	3	3
C9J9K3	40S ribosomal protein SA (Fragment) OS=Homo sapiens GN=RPS/	3	3	3	3
Q16555	Dihydropyrimidinase-related protein 2 OS=Homo sapiens GN=DPY	3	3	4	4
P62805	Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2 - [H4_I	6	6	6	6
P48047	ATP synthase subunit O, mitochondrial OS=Homo sapiens GN=AT	3	3	3	3
Q81V08	Phospholipase D3 OS=Homo sapiens GN=PLD3 PE=1 SV=1 - [PL	4	4	4	4
O75223	Gamma-glutamylcyclotransferase OS=Homo sapiens GN=GGCT PE	4	4	4	4
A0A0J9YXB3	Uncharacterized protein OS=Homo sapiens PE=4 SV=1 - [A0A0J9	4	4	4	4
A0A087WX12	IgGfC-binding protein OS=Homo sapiens GN=FCGBP PE=1 SV=1	5	5	5	5
G3V1A4	Cofilin 1 (Non-muscle), isoform CRA_a OS=Homo sapiens GN=CFL	3	3	3	3
P12273	Prolactin-inducible protein OS=Homo sapiens GN=PIP PE=1 SV=1	3	3	3	3
P05089	Arginase-1 OS=Homo sapiens GN=ARG1 PE=1 SV=2 - [ARG1_H	3	3	3	3
P20774	Mimcan OS=Homo sapiens GN=OGN PE=1 SV=1 - [MIME_HUM/	3	3	3	3
A0A075B6K4	HCG2043238 (Fragment) OS=Homo sapiens GN=IGLV3-10 PE=1	2	2	2	3
C9JGI3	Thymidine phosphorylase (Fragment) OS=Homo sapiens GN=TYM	4	4	4	5
G1TC33	Uncharacterized protein OS=Oryctolagus cuniculus GN=MYH10 PE	1	1	6	6
P19013	Keratin, type II cytoskeletal 4 OS=Homo sapiens GN=KRT4 PE=1	1	1	2	2
G1T277	Actin-related protein 2/3 complex subunit 3 OS=Oryctolagus cunic	2	2	2	3
A0A0A0MS15	Protein IGHV3-49 (Fragment) OS=Homo sapiens GN=IGHV3-49 P	2	2	3	3

H3B8MH2	Ras-related protein Rab-11A (Fragment) OS=Homo sapiens GN=R	3	3	3
E9PNW4	CD59 glycoprotein OS=Homo sapiens GN=CD59 PE=1 SV=1 - [E9	3	3	3
Q5SRP5	Apolipoprotein M OS=Homo sapiens GN=APOM PE=1 SV=1 - [Q5	3	3	3
A0A0C4DH32	Protein IGHV3-20 (Fragment) OS=Homo sapiens GN=IGHV3-20 PI	1	3	5
X6RJP6	Transgelin-2 (Fragment) OS=Homo sapiens GN=TAGLN2 PE=1 SV	5	5	7
P35579	Myosin-9 OS=Homo sapiens GN=MYH9 PE=1 SV=4 - [MYH9_HUN	2	5	5
C9J0E4	Cystatin-A OS=Homo sapiens GN=CSTA PE=1 SV=1 - [C9J0E4_HI	2	2	2
P09382	Galectin-1 OS=Homo sapiens GN=LGALS1 PE=1 SV=2 - [LEG1_HI	2	2	2
F8VZY9	Keratin, type I cytoskeletal 18 OS=Homo sapiens GN=KRT18 PE=	2	3	4
Q9UJR2	Cathepsin Z OS=Homo sapiens GN=CTSZ PE=1 SV=1 - [CATZ_HU	2	2	2
P01709	Immunoglobulin lambda variable 2-8 OS=Homo sapiens GN=IGLV	2	2	3
H0YD13	CD44 antigen OS=Homo sapiens GN=CD44 PE=1 SV=2 - [H0YD1	3	3	3
P02792	Ferritin light chain OS=Homo sapiens GN=FTL PE=1 SV=2 - [FRIL	2	2	2
H0YFA4	Cysteine-rich protein 2 (Fragment) OS=Homo sapiens GN=CRIP2	4	4	5
A0A0B4J1X8	Protein IGHV3-43 (Fragment) OS=Homo sapiens GN=IGHV3-43 PI	1	2	4
P81605	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2 - [DCD_HUMA	2	2	4
Q14956	Transmembrane glycoprotein NMB OS=Homo sapiens GN=GPNMB	2	2	4
A0A0C4DH31	Protein IGHV1-18 (Fragment) OS=Homo sapiens GN=IGHV1-18 PI	1	2	2
O14880	Microsomal glutathione S-transferase 3 OS=Homo sapiens GN=MC	2	2	2
I3L1H9	Zymogen granule protein 16 homolog B (Fragment) OS=Homo sap	2	2	2
K7ES69	Calponin-2 OS=Homo sapiens GN=CNN2 PE=1 SV=1 - [K7ES69_I	1	2	2
Q5T749	Keratinocyte proline-rich protein OS=Homo sapiens GN=KPRP PE=	4	4	4
A0A0C4DH29	Protein IGHV1-3 (Fragment) OS=Homo sapiens GN=IGHV1-3 PE=	2	2	2
Q16270	Insulin-like growth factor-binding protein 7 OS=Homo sapiens GN:	3	3	3
H0YNE9	Ras-related protein Rab-8B (Fragment) OS=Homo sapiens GN=RA	2	4	4
G1TL92	Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=2 - [2	2	2
A0A087WUQ6	Glutathione peroxidase OS=Homo sapiens GN=GPX1 PE=1 SV=1	2	2	2
A0A0B4J2D9	Protein IGKV1D-13 (Fragment) OS=Homo sapiens GN=IGKV1D-13	2	2	4
P51148	Ras-related protein Rab-5C OS=Homo sapiens GN=RAB5C PE=1 S	1	2	3
P01704	Immunoglobulin lambda variable 2-14 OS=Homo sapiens GN=IGL	1	1	3
Q08554	Desmocollin-1 OS=Homo sapiens GN=DSC1 PE=1 SV=2 - [DSC1	2	2	2
P46940	Ras GTPase-activating-like protein IQGAP1 OS=Homo sapiens GN:	4	4	4
V9GYE3	Apolipoprotein A-II OS=Homo sapiens GN=APOA2 PE=1 SV=1 - [2	2	2
Q99536	Synaptic vesicle membrane protein VAT-1 homolog OS=Homo sap	3	3	3
O75083	WD repeat-containing protein 1 OS=Homo sapiens GN=WDR1 PE=	3	3	3
P02748	Complement component C9 OS=Homo sapiens GN=C9 PE=1 SV=	2	2	3
P51911	Calponin-1 OS=Homo sapiens GN=CNN1 PE=1 SV=2 - [CNN1_HU	4	4	4
P35442	Thrombospondin-2 OS=Homo sapiens GN=THBS2 PE=1 SV=2 - [5	5	5
P15814	Immunoglobulin lambda-like polypeptide 1 OS=Homo sapiens GN=	2	2	3

I6L8B7	Fatty acid-binding protein, epidermal OS=Homo sapiens GN=FABP	2	2	3
P04217	Alpha-1B-glycoprotein OS=Homo sapiens GN=A1BG PE=1 SV=4	2	2	2
Q92765	Secreted frizzled-related protein 3 OS=Homo sapiens GN=FRZB PE	2	2	2
P20851	C4b-binding protein beta chain OS=Homo sapiens GN=C4BPB PE=	3	3	3
P31025	Lipocalin-1 OS=Homo sapiens GN=LCN1 PE=1 SV=1 - [LCN1_HUI	2	2	2
P59665	Neutrophil defensin 1 OS=Homo sapiens GN=DEFA1 PE=1 SV=1	2	2	2
P17813	Endoglin OS=Homo sapiens GN=ENG PE=1 SV=2 - [EGLN_HUMA]	2	2	2
Q14574	Desmocollin-3 OS=Homo sapiens GN=DSC3 PE=1 SV=3 - [DSC3_	3	3	3
P55058	Phospholipid transfer protein OS=Homo sapiens GN=PLTP PE=1 S	4	4	5
P30043	Flavin reductase (NADPH) OS=Homo sapiens GN=BLVRB PE=1 SV	2	2	2
A0A0B4J1V0	Protein IGHV3-15 (Fragment) OS=Homo sapiens GN=IGHV3-15 PI	1	1	2
P01699	Immunoglobulin lambda variable 1-44 OS=Homo sapiens GN=IGL	1	1	3
G1SRA9	Uncharacterized protein OS=Oryctolagus cuniculus GN=ILK PE=4	4	4	5
G1SQR6	Calponin OS=Oryctolagus cuniculus PE=3 SV=1 - [G1SQR6_RABIT	1	1	2
P16152	Carbonyl reductase [NADPH] 1 OS=Homo sapiens GN=CBR1 PE=	2	2	2
G1SPP0	Uncharacterized protein OS=Oryctolagus cuniculus GN=STOM PE=	2	2	2
A0A075B6R9	Protein IGV2D-24 (Fragment) OS=Homo sapiens GN=IGV2D-24	1	1	2
G1TWA7	Fibrinogen alpha chain OS=Oryctolagus cuniculus GN=FGA PE=4	1	1	2
A0A075B6I0	Protein IGLV8-61 (Fragment) OS=Homo sapiens GN=IGLV8-61 PE	2	2	2
M0R2L9	40S ribosomal protein S19 (Fragment) OS=Homo sapiens GN=RPS	2	2	2
P36955	Pigment epithelium-derived factor OS=Homo sapiens GN=SERPINI	3	3	3
A0A075B6J9	HCG1782423 (Fragment) OS=Homo sapiens GN=IGLV2-18 PE=1	2	2	2
P55083	Microfibril-associated glycoprotein 4 OS=Homo sapiens GN=MFAP	2	2	3
A0A0J9YY65	Alpha-2-antiplasmin (Fragment) OS=Homo sapiens GN=SERPINF2	2	2	2
Q6KB66	Keratin, type II cytoskeletal 80 OS=Homo sapiens GN=KRT80 PE=	2	2	2
A0A087WSY5	Carboxypeptidase B2 OS=Homo sapiens GN=CPB2 PE=1 SV=1 - [2	2	2
P00736	Complement C1r subcomponent OS=Homo sapiens GN=C1R PE=	2	2	2
G1SRF7	Uncharacterized protein OS=Oryctolagus cuniculus GN=HSPA9 PE	2	2	2
E7EWE8	Filamin-binding LIM protein 1 (Fragment) OS=Homo sapiens GN=	2	2	2
H9KV75	Alpha-actinin-1 OS=Homo sapiens GN=ACTN1 PE=1 SV=1 - [H9K	3	3	3
O14791	Apolipoprotein L1 OS=Homo sapiens GN=APOL1 PE=1 SV=5 - [AF	3	3	3
DGRBL5	Annexin OS=Homo sapiens GN=ANXA5 PE=1 SV=1 - [D6RBL5_HI	3	3	3
G1SGH0	Uncharacterized protein OS=Oryctolagus cuniculus GN=LOC10033	1	1	5
A0A0A0MR02	Voltage-dependent anion-selective channel protein 2 (Fragment) C	2	2	2
A0A0B4J2I3	60S ribosomal protein L30 OS=Homo sapiens GN=RPL30 PE=1 SV	2	2	2
A0A0B4J1Y8	Protein IGLV9-49 OS=Homo sapiens GN=IGLV9-49 PE=1 SV=1	3	3	3
G1U978	Uncharacterized protein OS=Oryctolagus cuniculus GN=CDC42 PE	1	1	2
F5H1C6	Fermitin family homolog 3 (Fragment) OS=Homo sapiens GN=FER	2	2	2
Q00610	Clastrin heavy chain 1 OS=Homo sapiens GN=CLTC PE=1 SV=5 -	7	7	7

G1T3Y8	Uncharacterized protein OS=Oryctolagus cuniculus GN=HSPD1 PE	2	2	2
A0A075B7D0	Protein IGHV1OR15-1 (Fragment) OS=Homo sapiens GN=IGHV1C	1	1	2
P21796	Voltage-dependent anion-selective channel protein 1 OS=Homo sa	2	2	2
A0A140T958	HLA class II histocompatibility antigen, DRB1-16 beta chain (Fragr	1	1	1
A0A075B7D8	Protein IGHV3OR15-7 (Fragment) OS=Homo sapiens GN=IGHV3C	1	1	2
G1U0D0	Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=1 - [1	1	2
E9PBC5	Plasma kallikrein OS=Homo sapiens GN=KLKB1 PE=1 SV=2 - [E9f	2	2	2
A0A0G2JQM1	AP-2 complex subunit alpha-2 (Fragment) OS=Homo sapiens GN=	2	2	2
Q16853	Membrane primary amine oxidase OS=Homo sapiens GN=AOC3 P	6	6	6
P01593	Immunoglobulin kappa variable 1D-33 OS=Homo sapiens GN=IGK	1	1	3
Q5D862	Flaggrin-2 OS=Homo sapiens GN=FLG2 PE=1 SV=1 - [FLA2_HUI	3	3	3
Q9NZN4	EH domain-containing protein 2 OS=Homo sapiens GN=EHD2 PE=	3	3	3
Q08188	Protein-glutamine gamma-glutamyltransferase E OS=Homo sapien	1	1	1
H7C106	Reticulon (Fragment) OS=Homo sapiens GN=RTN4 PE=1 SV=1 -	1	1	1
G1U9R7	Uncharacterized protein OS=Oryctolagus cuniculus GN=COL4A1 P	1	1	1
P01703	Immunoglobulin lambda variable 1-40 OS=Homo sapiens GN=IGL	1	1	1
A6NNI4	Tetraspanin OS=Homo sapiens GN=CD9 PE=1 SV=1 - [A6NNI4_H	2	2	2
P10643	Complement component C7 OS=Homo sapiens GN=C7 PE=1 SV=	4	4	4
H3BSW3	Adenine phosphoribosyltransferase OS=Homo sapiens GN=APRT F	2	2	2
Q6NZI2	Polymerase I and transcript release factor OS=Homo sapiens GN=	1	1	1
P63167	Dynein light chain 1, cytoplasmic OS=Homo sapiens GN=DYNLL1	1	1	1
B1AN99	Trypsin-3 (Fragment) OS=Homo sapiens GN=PRSS3 PE=1 SV=8 -	1	1	1
E9PND2	Cysteine and glycine-rich protein 1 (Fragment) OS=Homo sapiens	4	4	5
F8VUA5	Ras-related protein Rab-5B (Fragment) OS=Homo sapiens GN=RA	1	2	2
P04430	Immunoglobulin kappa variable 1-16 OS=Homo sapiens GN=IGKV	2	2	2
P20160	Azurocidin OS=Homo sapiens GN=AZU1 PE=1 SV=3 - [CAP7_HUI	1	1	1
A0A0A0MT36	Protein IGKV6D-21 (Fragment) OS=Homo sapiens GN=IGKV6D-21	2	2	2
Q99714	3-hydroxyacyl-CoA dehydrogenase type-2 OS=Homo sapiens GN=	3	3	3
F8WE65	Peptidyl-prolyl cis-trans isomerase OS=Homo sapiens GN=PPIA PE	2	2	2
H7C3T4	Peroxisome oxidin-4 (Fragment) OS=Homo sapiens GN=PRDX4 PE=1	1	3	3
F5H0V9	Germ cell-specific gene 1 protein (Fragment) OS=Homo sapiens G	1	1	1
P05109	Protein S100-A8 OS=Homo sapiens GN=S100A8 PE=1 SV=1 - [S1	1	1	1
P30050	60S ribosomal protein L12 OS=Homo sapiens GN=RPL12 PE=1 SV	2	2	2
C9JC71	Low affinity immunoglobulin gamma Fc region receptor III-A (Frag	2	2	2
G1TV31	Actin-related protein 2/3 complex subunit 5 OS=Oryctolagus cunic	1	1	1
F2Z2S8	40S ribosomal protein S3 OS=Homo sapiens GN=RPS3 PE=1 SV=	2	2	2
F8VPE8	60S acidic ribosomal protein P0 (Fragment) OS=Homo sapiens GN	2	2	2
A0A0G2J3C0	Protein IGLV5-45 (Fragment) OS=Homo sapiens GN=IGLV5-45 PE	2	2	2
P02765	Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=	1	1	1

E9PJL7	Alpha-crystallin B chain (Fragment) OS=Homo sapiens GN=CRYAB	2	2	2
D6RD17	Immunoglobulin J chain (Fragment) OS=Homo sapiens GN=JCHAI	1	1	2
P00740	Coagulation factor IX OS=Homo sapiens GN=F9 PE=1 SV=2 - [FA	1	1	1
Q9UGM3	Deleted in malignant brain tumors 1 protein OS=Homo sapiens GN	1	1	1
P11279	Lysosome-associated membrane glycoprotein 1 OS=Homo sapiens	1	1	1
P32119	Peroxiredoxin-2 OS=Homo sapiens GN=PRDX2 PE=1 SV=5 - [PRC	2	3	3
F5H6Q2	Polyubiquitin-C (Fragment) OS=Homo sapiens GN=UBC PE=1 SV=	3	3	3
B1AK87	Capping protein (Actin filament) muscle Z-line, beta, isoform CRA_	2	2	2
X6RLJ0	Complement C1q subcomponent subunit A (Fragment) OS=Homo	1	1	1
P01687	Ig kappa chain V region BS-5 OS=Oryctolagus cuniculus PE=1 SV=	2	2	2
P30837	Aldehyde dehydrogenase X, mitochondrial OS=Homo sapiens GN=	2	2	2
A0A0C4DH34	Protein IGHV4-28 (Fragment) OS=Homo sapiens GN=IGHV4-28 PI	1	1	1
O14773	Tripeptidyl-peptidase 1 OS=Homo sapiens GN=TPP1 PE=1 SV=2	2	2	2
C9JSN8	High affinity immunoglobulin gamma Fc receptor I (Fragment) OS=	1	1	1
F5H7Z2	Microfibrillar-associated protein 5 OS=Homo sapiens GN=MFAP5 P	1	1	1
Q28668	Collagen alpha-2(I) chain (Fragment) OS=Oryctolagus cuniculus G	1	1	1
Q14568	Heat shock protein HSP 90-alpha A2 OS=Homo sapiens GN=HSP9	1	1	1
Q8WUJM4	Programmed cell death 6-interacting protein OS=Homo sapiens GN	1	1	1
P63092	Guanine nucleotide-binding protein G(s) subunit alpha isoforms sh	1	2	2
K7ER19	Apolipoprotein C-I (Fragment) OS=Homo sapiens GN=APOC1 PE=	1	1	1
P61106	Ras-related protein Rab-14 OS=Homo sapiens GN=RAB14 PE=1 S	1	2	2
P55001	Microfibrillar-associated protein 2 OS=Homo sapiens GN=MFAP2 P	1	1	1
H0YL80	Tropomyosin alpha-1 chain (Fragment) OS=Homo sapiens GN=TP	1	1	1
P63104	14-3-3 protein zeta/delta OS=Homo sapiens GN=YWHAZ PE=1 SV	2	2	2
A0A075B619	Protein IGLV7-46 (Fragment) OS=Homo sapiens GN=IGLV7-46 PE	1	1	1
A0A0U1RQT9	Synaptophysin-like protein 1 (Fragment) OS=Homo sapiens GN=S	1	1	1
K7EP04	Heat shock protein beta-6 OS=Homo sapiens GN=HSPB6 PE=1 SV	1	1	1
Q08722	Leukocyte surface antigen CD47 OS=Homo sapiens GN=CD47 PE=	1	1	1
H3BRU6	Poly(rC)-binding protein 2 (Fragment) OS=Homo sapiens GN=PCB	2	3	3
Q15113	Procollagen C-endopeptidase enhancer 1 OS=Homo sapiens GN=F	1	1	1
P05546	Heparin cofactor 2 OS=Homo sapiens GN=SERPIND1 PE=1 SV=3	4	4	4
E7EQJ3	HLA class II histocompatibility antigen gamma chain OS=Homo sa	1	1	1
P54289	Voltage-dependent calcium channel subunit alpha-2/delta-1 OS=H	2	2	2
C9K028	Nucleoside diphosphate kinase A (Fragment) OS=Homo sapiens G	1	1	1
A0A0J9YX35	Uncharacterized protein (Fragment) OS=Homo sapiens PE=1 SV=	2	2	2
A0A087X0K0	Collagen alpha-1(XV) chain OS=Homo sapiens GN=COL15A1 PE=	2	2	2
C9JUG7	F-actin-capping protein subunit alpha-2 OS=Homo sapiens GN=CA	1	1	1
A0A0B4J1U7	Protein IGHV6-1 (Fragment) OS=Homo sapiens GN=IGHV6-1 PE=	1	1	1
G1TME7	Uncharacterized protein OS=Oryctolagus cuniculus GN=CSTB PE=	1	1	1

P04083	Annexin A1 OS=Homo sapiens GN=ANXA1 PE=1 SV=2 - [ANXA1_	3	3	3
R4GN08	Actin-related protein 2/3 complex subunit 4 (Fragment) OS=Homo	1	1	1
Q5H9B4	Metalloproteinase inhibitor 1 (Fragment) OS=Homo sapiens GN=T	1	1	1
P01601	Immunoglobulin kappa variable 1D-16 (Fragment) OS=Homo sapi	1	1	1
B1AHF3	NADH-cytochrome b5 reductase 3 (Fragment) OS=Homo sapiens (1	1	1
F8VXZ8	Decorin (Fragment) OS=Homo sapiens GN=DCN PE=1 SV=1 - [F8	1	1	1
G1TE13	Uncharacterized protein OS=Oryctolagus cuniculus GN=RHOG PE=	1	2	2
P25788	Proteasome subunit alpha type-3 OS=Homo sapiens GN=PSMA3 P	1	1	1
P62318	Small nuclear ribonucleoprotein Sm D3 OS=Homo sapiens GN=SN	1	1	1
H0YEL7	Acetyl-CoA acetyltransferase, mitochondrial (Fragment) OS=Homo	1	1	1
H3BMQ8	Fructose-bisphosphate aldolase A (Fragment) OS=Homo sapiens C	2	2	2
G1T235	Proteasome subunit beta type OS=Oryctolagus cuniculus GN=PSM	1	1	1
Q15582	Transforming growth factor-beta-induced protein ig-h3 OS=Homo	2	2	2
Q8TCE1	Antithrombin-III OS=Homo sapiens GN=SERPINC1 PE=1 SV=1 - [1	1	1
P02787	Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=3 - [TRFE_HU	3	3	3
E9PQR3	Ferritin OS=Homo sapiens GN=FTH1 PE=1 SV=1 - [E9PQR3_HUM	1	1	1
H0YMQ8	Protein-glutamine gamma-glutamyltransferase K (Fragment) OS=H	1	1	1
G1TY53	Ribosomal protein OS=Oryctolagus cuniculus PE=3 SV=1 - [G1TY:	1	1	1
P05556	Integrin beta-1 OS=Homo sapiens GN=ITGB1 PE=1 SV=2 - [ITB1	2	2	2
G3V5I3	Alpha-1-antichymotrypsin (Fragment) OS=Homo sapiens GN=SER	1	1	1
U3KM64	Uncharacterized protein OS=Oryctolagus cuniculus GN=CLTA PE=	1	1	1
M0QZN2	40S ribosomal protein S5 OS=Homo sapiens GN=RPS5 PE=1 SV=	1	1	1
P06702	Protein S100-A9 OS=Homo sapiens GN=S100A9 PE=1 SV=1 - [S1	1	1	1
A0A087WYS6	Proteasome (Prosome, macropain) subunit, alpha type, 8, isoform	2	2	2
Q14764	Major vault protein OS=Homo sapiens GN=MVP PE=1 SV=4 - [MV	3	3	3
A0A0A0MT20	EMILIN-1 (Fragment) OS=Homo sapiens GN=EMILIN1 PE=1 SV=	1	1	1
P00742	Coagulation factor X OS=Homo sapiens GN=F10 PE=1 SV=2 - [FA	1	1	1
P07858	Cathepsin B OS=Homo sapiens GN=CTSB PE=1 SV=3 - [CATB_HU	3	3	3
P12318	Low affinity immunoglobulin gamma Fc region receptor II-a OS=H	1	1	1
G1TF67	Coronin OS=Oryctolagus cuniculus GN=CORO1A PE=3 SV=1 - [G1	1	1	1
M0QX76	40S ribosomal protein S16 (Fragment) OS=Homo sapiens GN=RPc	1	1	1
B1AHL2	Fibulin-1 OS=Homo sapiens GN=FBLN1 PE=1 SV=1 - [B1AHL2_HU	1	1	1
G1T7A8	Uncharacterized protein OS=Oryctolagus cuniculus GN=CPQ PE=4	1	1	1
Q07960	Rho GTPase-activating protein 1 OS=Homo sapiens GN=ARHGAP1	1	1	1
G1T7F1	Histone H2B OS=Oryctolagus cuniculus GN=LOC100348835 PE=3	3	3	3
A0A0B4J1V2	Protein IGHV2-26 (Fragment) OS=Homo sapiens GN=IGHV2-26 PI	1	1	1
B1AKG0	Complement factor H-related protein 1 OS=Homo sapiens GN=CFI	1	1	1
G1SFP0	Uncharacterized protein OS=Oryctolagus cuniculus GN=LAP3 PE=	2	2	2
G1SZ15	Uncharacterized protein OS=Oryctolagus cuniculus PE=1 SV=1 - [2	2	2

P29622	Kallistatin OS=Homo sapiens GN=SERPINA4 PE=1 SV=3 - [KAIN_	1	1	1
G3V511	Latent-transforming growth factor beta-binding protein 2 OS=Horr	1	1	1
G1T319	Annexin OS=Oryctolagus cuniculus GN=ANXA7 PE=3 SV=1 - [G1T	1	1	1
G1U4G9	Chloride intracellular channel protein OS=Oryctolagus cuniculus G	1	1	1
P01847	Ig lambda chain C region OS=Oryctolagus cuniculus PE=1 SV=2 -	1	1	1
A0A0C4DH36	Protein IGHV3-38 (Fragment) OS=Homo sapiens GN=IGHV3-38 P	1	1	1
H0Y4U1	Tensin-1 OS=Homo sapiens GN=TNS1 PE=1 SV=1 - [H0Y4U1_HU	1	1	1
G1TVL1	Elongation factor 1-gamma OS=Oryctolagus cuniculus GN=EEF1G	1	1	1
P80108	Phosphatidylinositol-glycan-specific phospholipase D OS=Homo sa	1	1	1
F8W7S5	Ribosome-binding protein 1 OS=Homo sapiens GN=RRBP1 PE=1 S	1	1	1
P60903	Protein S100-A10 OS=Homo sapiens GN=S100A10 PE=1 SV=2 - [1	1	2
D6RAA6	Transmembrane protein 33 (Fragment) OS=Homo sapiens GN=TM	1	1	1
P11021	78 kDa glucose-regulated protein OS=Homo sapiens GN=HSPA5 P	1	2	2
P28070	Proteasome subunit beta type-4 OS=Homo sapiens GN=PSMB4 PE	1	1	1
G1SND0	Sulfhydryl oxidase OS=Oryctolagus cuniculus GN=QSOX1 PE=4 S	1	1	1
P18428	Lipopolysaccharide-binding protein OS=Homo sapiens GN=LBP PE	1	1	1
J3KRL0	Unconventional myosin-1d OS=Homo sapiens GN=MYO1D PE=1 S	1	1	1
P53680	AP-2 complex subunit sigma OS=Homo sapiens GN=AP2S1 PE=1	1	1	1
H7C463	MICOS complex subunit MIC60 (Fragment) OS=Homo sapiens GN	1	1	1
J3KTJ1	Myosin regulatory light chain 12A (Fragment) OS=Homo sapiens G	1	1	1
Q9UHG3	Prenylcysteine oxidase 1 OS=Homo sapiens GN=PCYOX1 PE=1 SV	2	2	2
G1TFE8	Uncharacterized protein OS=Oryctolagus cuniculus GN=RPS26 PE	1	1	1
G1SIQ6	Uncharacterized protein OS=Oryctolagus cuniculus GN=CPNE4 PE	1	1	1
P08311	Cathepsin G OS=Homo sapiens GN=CTSG PE=1 SV=2 - [CATG_H	1	1	1
P10253	Lysosomal alpha-glucosidase OS=Homo sapiens GN=GAA PE=1 S	1	1	1
G1T6P5	Uncharacterized protein OS=Oryctolagus cuniculus GN=ARL8B PE	1	1	1
G1T6S2	Uncharacterized protein OS=Oryctolagus cuniculus GN=PGM5 PE=	1	1	1
A0A0A0MSQ0	Plastin-3 OS=Homo sapiens GN=PLS3 PE=1 SV=1 - [A0A0A0MSQ	2	2	2
P08603	Complement factor H OS=Homo sapiens GN=CFH PE=1 SV=4 - [C	2	2	2
P21589	5'-nucleotidase OS=Homo sapiens GN=NT5E PE=1 SV=1 - [5NTD	3	3	4
P40261	Nicotinamide N-methyltransferase OS=Homo sapiens GN=NNMT P	1	1	1
Q5VVC9	60S ribosomal protein L11 (Fragment) OS=Homo sapiens GN=RPL	1	1	1
P41218	Myeloid cell nuclear differentiation antigen OS=Homo sapiens GN=	3	3	3
J3QS7	Myeloperoxidase (Fragment) OS=Homo sapiens GN=MPO PE=1 S	1	1	1
G1TGC7	Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=2 - [1	1	1
Q5T750	Skin-specific protein 32 OS=Homo sapiens GN=XP32 PE=1 SV=1	1	1	1
A0A1B0GV06	Acid ceramidase (Fragment) OS=Homo sapiens GN=ASAH1 PE=1	1	1	1
B4DHN5	Syntenin-1 OS=Homo sapiens GN=SDCBP PE=1 SV=1 - [B4DHN5	1	1	1
G1TKA1	Uncharacterized protein OS=Oryctolagus cuniculus PE=4 SV=1 - [1	1	2

D6RBJ7	Vitamin D-binding protein OS=Homo sapiens GN=GC PE=1 SV=1	2	2	2
F5H7V9	Tenascin OS=Homo sapiens GN=TNC PE=1 SV=1 - [F5H7V9_HUN	2	2	2
P26038	Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 - [MOES_HUMAN	2	2	2

2ème partie: Criblage *in silico* d'anticorps humains

I Séquençage haut débit de phage-scFv issus d'une sélection *in vivo*

I.1 Introduction

Dans le but de rechercher de nouvelles cibles dans le cadre de l'athérosclérose, des anticorps ont été sélectionnés *in vivo* puis caractérisés *in vitro* par la technique de cytométrie de flux, développée pendant ces trois années de thèse. Cette étape de caractérisation a été suivie, comme dans beaucoup d'études de criblage, d'un séquençage Sanger pour déterminer la séquence de l'anticorps d'intérêt. Cet enchaînement de procédés m'a permis d'aboutir à la sélection de deux cents anticorps ainsi que la découverte d'une cible : la galectine 3. Cependant, parmi des centaines de milliers de phage-scFv sélectionnés *in vivo*, seuls 2000 clones ont pu être caractérisés *in vitro*. Cette caractérisation étant donc extrêmement longue, de nouvelles techniques apparaissent depuis ces dernières années comme méthodes de criblage subsidiaires. Ainsi, le séquençage haut débit (NGS (Next generation sequencing)) d'immunoglobulines (Ig) a pris une place prépondérante avec la possibilité de séquencer des centaines de milliers de molécules en un temps court. Les premières techniques de séquençage ont permis de séquencer de manière séparée un gène V_H d'un gène V_L . Prenons en exemple l'étude de Reddy *et al* qui a mis en exergue et développé des anticorps spécifiques de trois cibles grâce à l'étude des fréquences d'un gène V_H et d'un gène V_L provenant de lymphocytes B de souris immunisées sans construction de banques d'anticorps⁴². Ainsi, le gène V_H le plus représenté a été fusionné au gène V_L le plus représenté. Après construction, le scFv a été produit sous format soluble pour être testé *in vitro* pour sa capacité à reconnaître sa cible et également mesurer son affinité. Lors de la sélection de banques peptidiques ou d'anticorps, cela devient encore plus complexe si l'on veut tenir compte du facteur « enrichissement ». Comme l'a souligné Michael F. Naso très justement, lors d'un criblage *in vitro*, il est difficile d'avoir un aperçu global de cet enrichissement mais également de l'unicité des anticorps avant les étapes finales d'un criblage, qui sont la caractérisation *in vitro* suivie d'un séquençage par Sanger⁴³. En 2012, W. Matochko a montré qu'il était possible par séquençage Illumina de mettre en exergue des clones sous-représentés dans une banque de peptides. Cette étude a permis de mettre en évidence un biais important de la sélection *in vivo* puis *in vitro*. En effet, les résultats ont suggéré que les criblages basés sur l'amplification répétée (différents tours de sélection) faisant perdre en diversité et le séquençage à petite échelle par Sanger ne permettent d'identifier que quelques clones parmi des milliers de clones potentiellement intéressants. Ainsi, cette approche par séquençage Illumina pourrait également être utilisée à des fins de criblage *in silico*, qui peut préserver la diversité des clones phage-scFv et identifier des ligands non sélectionnés dans les études de criblage *in vitro*⁴⁴. En parallèle, U. Ravn a démontré que la technique Illumina pouvait également être utilisée pour analyser l'enrichissement de scFv contre deux antigènes⁴⁵. Après l'analyse des régions CDR3 de chaînes lourdes les plus représentées, des amorces ont pu être dessinées afin de récupérer, au sein du pool des séquences, les clones les plus enrichis du troisième tour de sélection.

Ainsi, ces études ont montré qu'il était possible d'accéder à l'enrichissement par des méthodes de séquençage haut débit. Cependant, aucune ne fait état de l'utilisation de ces techniques pour un séquençage d'un scFv entier sur plus de 800pb. En effet, la plupart des techniques utilisées jusqu'à aujourd'hui se limite à l'analyse d'un maximum de 400pb. Le développement par Pacific Biosciences d'outils de séquençage pouvant aller jusqu'à l'analyse de fragments d'une taille supérieure à 25kpb pourrait permettre d'obtenir, en une seule analyse, les deux domaines V d'un scFv (taille >800pb) qui sont complexes et très diversifiés. En effet, une diversité très importante est générée au niveau des domaines V d'une immunoglobuline, en particulier dans les régions hypervariables et notamment dans le CDR3 durant le réarrangement de l'ADN. Un autre niveau de diversité est également apporté lors de

la maturation de l'affinité appelée « hypermutation somatique ». P. Mondon a mimé *in vitro* ce procédé grâce à l'utilisation de polymérase dites « mutases » qui ont généré des mutations aléatoires lors de la construction de la banque que j'ai utilisée pendant ma thèse.

Par ailleurs, Marie Paule LEFRANC a développé, en 1989, un système d'information des immunoglobulines, regroupant tous les fondements de l'immunologie, des mécanismes impliqués dans la diversité des immunoglobulines mais également permettant l'analyse des clonotypes. Suite à une interaction anticorps-antigène spécifique, la cellule B prolifère et génère des clones engagés dans la sélection *in vivo*, la maturation de l'affinité et regroupés dans un même « clonotype ». Ce « clonotype » est défini par les mêmes gènes et allèles V, D et J et une même jonction CDR3.

La problématique de l'étude ici a donc été d'établir si (1) l'utilisation de la technique PacBio RS II pouvait apporter des données en grande quantité et de grande qualité ainsi que si (2) l'utilisation de la plateforme IMGT/V-QUEST et notamment sa fonctionnalité « scFv » pouvait rendre compte de possibles mutations générées lors de la création de la banque mais également de l'enrichissement des clones qui s'est effectuée durant la sélection *in vivo*.

I.2 Séquençage d'anticorps issus d'une sélection par phage display *in vivo*

Après la sélection par phage display *in vivo* d'une banque hyperdiversifiée de fragments d'anticorps (sélection expliquée dans la section « résultats » en partie 1, criblage *in vitro*), plus de 10^7 phages ont été récupérés en plusieurs fractions. Nous avons ici, dans l'étude pilote, étudié l'une des fractions issues du tour 3 : E2F1 P3 (diversité de $3,51 \cdot 10^5$) (**Tableau 5**). Afin de séquencer ces clones, ces fragments scFv ont été amplifiés par PCR (Polymerase Chain Reaction).

Tableau 5. Quantité de phage-scFv élués après les différents tours de sélection (P1 à P3) dans les différentes fractions. La quantité de phage-scFv est ici calculée en cfu après étalement sur les boîtes de pétri.

	P1	P2	P3
E1F1	1,05E+07	4,95E+04	7,63E+05
E2F1	1,79E+06	1,58E+05	3,51E+05
E1F2	3,31E+06	3,66E+04	2,8E+06
E2F2	6,79E+06	3,35E+05	9,80E+06
E1F3	3,66E+04	2,10E+03	6,25E+05
E2F3	2,86E+05		4,40E+05

I.2.1 Amplification par PCR de clones scFv

Afin d'amplifier les séquences d'ADN codant pour les scFv, des amorces ont été dessinées sur le vecteur ayant servi au clonage de la banque (**Figure 56**). Le schéma ci-après donne la représentation de l'amplification.

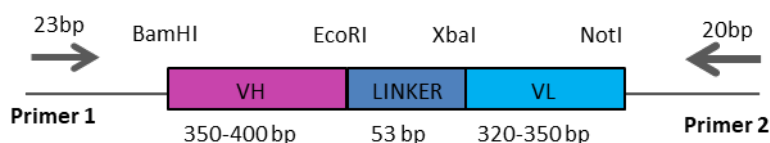


Figure 56 Schématisation de l'amplification PCR effectuée avec placement des amorces sur le vecteur

Afin d'obtenir des produits PCR purs, plusieurs polymérase, plusieurs nombres de cycles et plusieurs dilutions de plasmide ont été essayés avant de définir les conditions idéales permettant d'obtenir des amplicons de 1000pb de très bonne qualité (**Figure 57**).

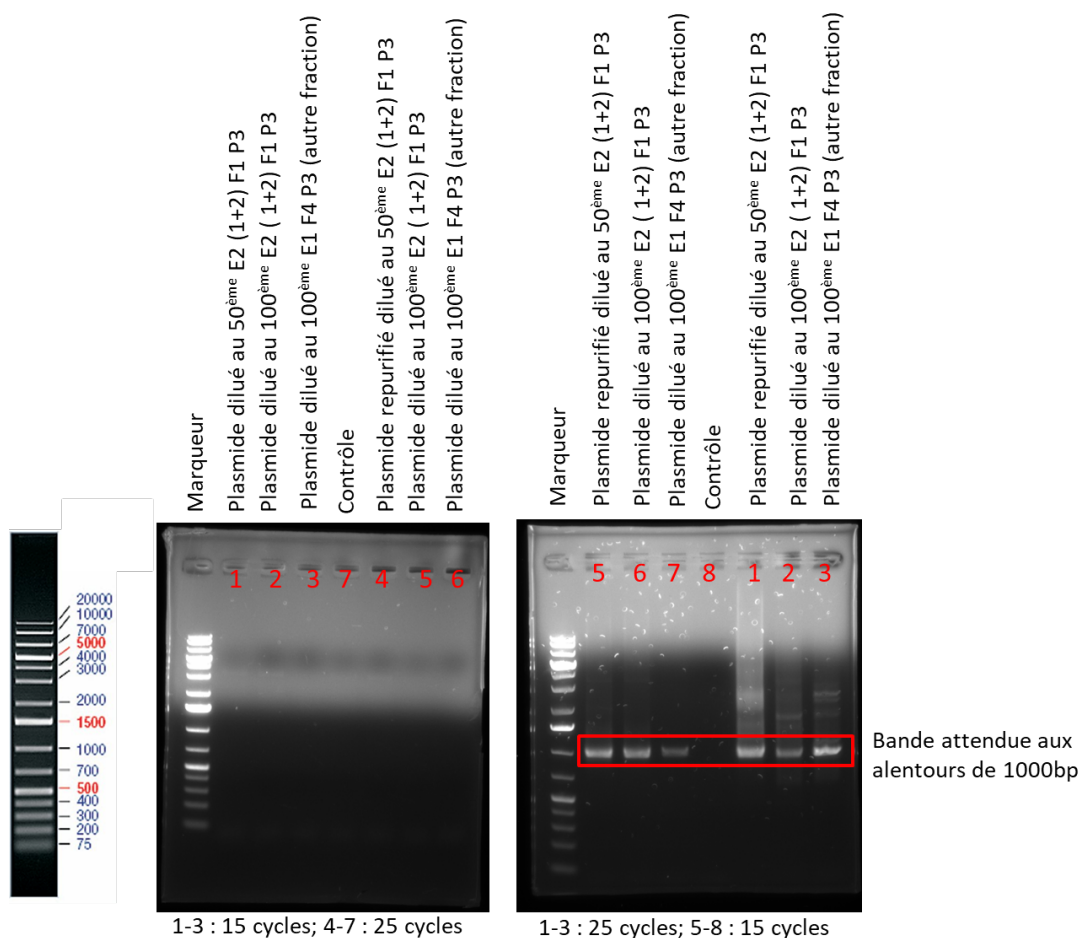


Figure 57 Résultats PCR des essais de différentes amorces, nombres de cycles et dilutions. Le gel 1 concerne l'amplification PCR avec la Taq polymérase. Le gel 2 concerne l'amplification avec la Kapa Hifi Hot Start polymérase.

Au vu des résultats, la Kapa Hifi Hot Start polymérase a été choisie pour le projet ainsi qu'un nombre de cycles équivalent à 15 et une dilution du plasmide au 100^{ème} (apparition de bandes parasites avec une amplification de 25 cycles) (**Figure 57**). Ainsi, quatre PCR ont été effectuées pour une comparaison des données. Le résultat de l'amplification est présenté ci-dessous (**Figure 58**).

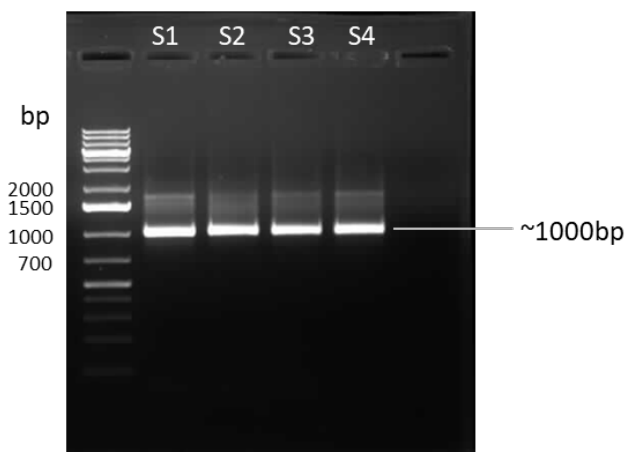


Figure 58 Résultat de l'amplification PCR des 4 produits avec l'utilisation de la polymérase Kapa Hifi Hot Start et 15 cycles.

Les produits PCR obtenus après cette amplification ont été testés par Bioanalyzer pour valider la pureté du produit puis séquencés par la méthode PacBio RS II (**Figure 59**).

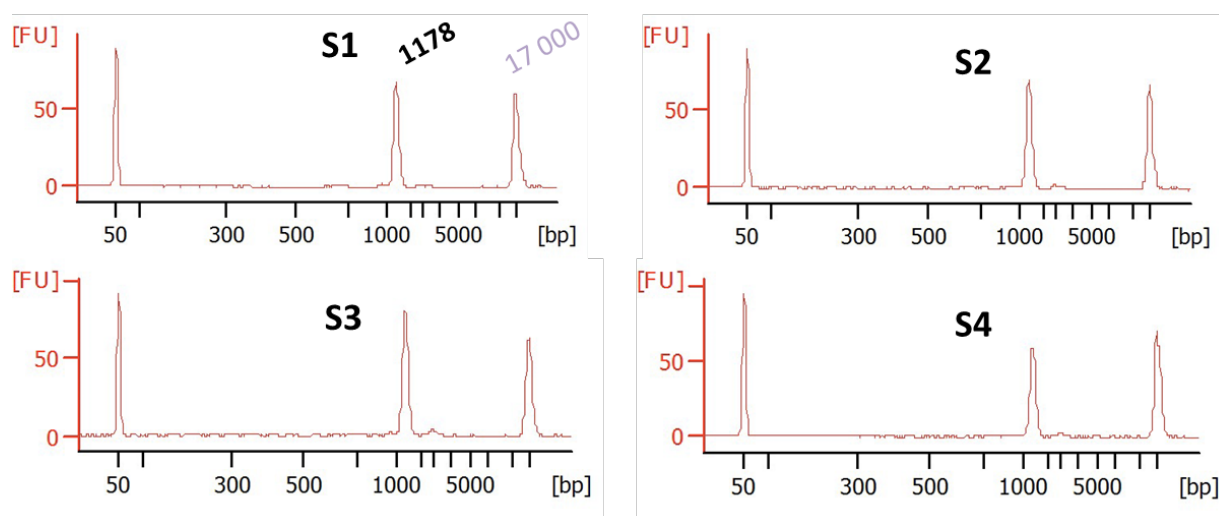


Figure 59 Résultat du Bioanalyzer sur les 4 produits PCR envoyés à Pacific Biosciences (Figure obtenues par nos collaborateurs chez Pacific Biosciences aux Etats-Unis)

I.2.2 Préparation et séquençage des scFv amplifiés

Le schéma ci-dessous fait état des différentes étapes de préparation des brins d'ADN ainsi que du processus de séquençage (**Figure 60**). Les brins d'ADN, une fois circularisés grâce à l'ajout d'adaptateurs suivi de la fixation des amorces et de la polymérase, sont chargés sur des puces dites « SMRT (Single Molecule Real Time) » comportant des milliers de puits. Il est à noter qu'un seul brin d'ADN peut être chargé au fond d'un puits. La circularisation permet de pouvoir faire une lecture du brin en « rolling circle ». En effet, la technique PacBio permet de lire environ 25kpb. Un scFv faisant environ 800pb (et 1000pb si on inclue les extrémités 5' et 3' comptabilisant les régions flanquantes du vecteur et les primers utilisés, voir **Figure 56**), il est possible de lire la séquence d'intérêt plus d'une vingtaine de fois. Cette multi lecture permet ensuite, par comparaison avec des logiciels intrinsèques à ces plateformes, nommés filtres CCS (Circular consensus sequence), de pouvoir déterminer une séquence dite « consensus », désignant la séquence la plus probable du brin séquencé. Ainsi, chaque brin d'ADN a été séquencé. Les milliers de données ont été récupérés puis traités par le logiciel IMG/HighV-QUEST notamment avec la fonctionnalité « scFv ».

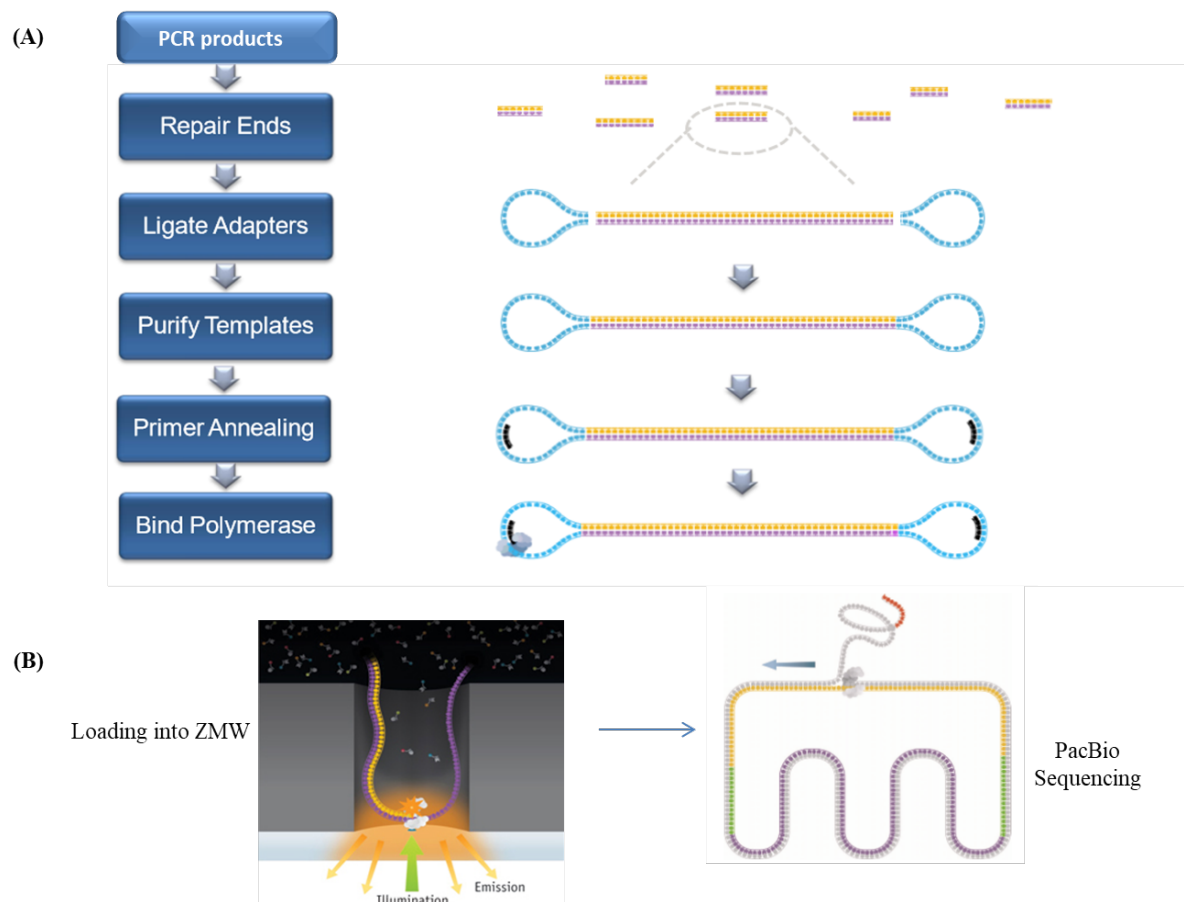


Figure 60 Schématisation de la préparation des brins d'ADN et processus de séquençage. (A) préparation de la librairie. Représentation des différentes étapes conduisant à l'obtention de brins d'ADN circularisés prêts à être chargés sur les SMRT. (B) Processus et schématisation du séquençage (Extrait du Site <http://pacific-biosciences.com>)

I.3 Résultats et discussion

L'ensemble des résultats sont décrits de manière détaillée dans l'article intitulé « PacBio sequencing and IMGT/HighV-QUEST analysis of full-length scFv from an *in vivo* selected phage-display combinatorial library » soumis dans la revue *Frontiers In Immunology* et dont je suis l'auteur principal.

Le séquençage des scFv comprenant les deux domaines V reliés entre eux par une courte séquence peptidique est devenu un réel besoin pour faire avancer les méthodes de criblage *in silico*. Ainsi nous avons montré qu'il était possible, grâce à l'utilisation de la lecture en « rolling circle » sur la plateforme PacBio RS II, d'obtenir des séquences de haute qualité pouvant être analysées en tant que scFv sur l'outil IMGT/HighV-QUEST. Grâce à cet outil, nous avons pu identifier 346 934 scFv. De manière à évaluer la qualité de ces « reads », la séquence du clone P3, anticorps favori dont la cible a été déterminée *in vitro*, a été comparée et recherchée dans les données de séquençage haut débit. Soixante reads ont pu être identifiés avec 100% d'identité à la séquence de P3 ainsi que 25 reads supplémentaires montrant des mutations ponctuelles sur la séquence. Parmi ces 25 reads, 15 d'entre eux ont été clairement identifiés comme appartenant au même clonotype (tenant compte des germes des gènes V, D et J ainsi que du CDR3) que le clone P3. De manière intéressante, nous avons pu remarquer

que le clone P3 faisait partie du top100 des clonotypes les plus représentés parmi les données sorties du séquençage haut débit.

Ceci a mené à deux conclusions : (1) l'utilisation de la technique PacBio suivie d'une analyse par la database IMGT peut être utilisée pour l'identification et la caractérisation complète de scFv mais également (2) ces deux techniques permettent d'identifier des « hits » de séquences préférentiellement enrichies au cours du processus de sélection *in vivo*.

Ainsi, le criblage *in silico* apparaît comme une nouvelle approche de criblage d'anticorps issus de sélection *in vivo* contre une multitude d'antigènes. Après une analyse approfondie des séquences et de leur enrichissement, les associations VH-VL les plus représentés pourront être synthétisées pour être ensuite testées *in vitro* pour la détermination des cibles.

Séquençage haut débit (NGS) et analyse
de scFv issus d'une sélection *in vivo*
(Article)

1 **PacBio sequencing and IMGT/HighV-QUEST analysis of full-length**
2 **scFv from an *in vivo* selected phage-display combinatorial library**

3 **Audrey Hemadou¹, Véronique Giudicelli², Melissa Smith^{3,6}, Marie-Paule Lefranc^{2*}, Patrice**
4 **Duroux², Sofia Kossida², Cheryl Heiner³, N. Lance Hepler³, John Kuijpers³, Alexis Groppi⁴,**
5 **Jonas Korlach³, Philippe Mondon⁵, Florence Ottonnes¹, Marie-Josée Jacobin-Valat¹, Jeanny**
6 **Laroche-Traineau^{1†} and Gisèle Clofent-Sanchez^{1†*}**

7 [†]Equivalent position of authors

8 ¹ CRMSB, UMR 5536, CNRS, Bordeaux, France

9 ² IMGT®, The international ImMunoGeneTics information system®, Laboratoire d'ImmunoGénétique
10 Moléculaire, LIGM, Institut de Génétique Humaine, IGH, UMR 9002, CNRS, Montpellier University,
11 Montpellier, France

12 ³ Pacific Biosciences, Menlo Park, California, USA

13 ⁴ Université de Bordeaux, Centre de Bioinformatique de Bordeaux (CBiB), Bordeaux, France

14 ⁵ LFB Biotechnologies, Lille, France

15 ⁶ Present address: Department of Genetics and Genomic Sciences, Icahn School of Medicine Mount
16 Sinai, New York, USA

17

18

19

20

21

22

23 *** Correspondence:**

24 Marie-Paule Lefranc

25 Marie-Paule.Lefranc@igh.cnrs.fr

26 Gisèle Clofent-Sanchez

27 gisele.clofent-sanchez@rmsb.u-bordeaux2.fr

28

29

30

31 **Running title:** Pacbio scFv sequencing and IMGT/HighV-QUEST analysis

32 **Keywords:** human antibody, IMGT/HighV-QUEST, immunoinformatics, immunoglobulin, PacBio
33 sequencing, phage combinatorial library, single chain scFv, next generation sequencing (NGS)

34

35

36 **Abstract**

37 Phage-display selection of immunoglobulin (IG) or antibody single chain Fragment variable (scFv)
38 from combinatorial libraries is widely used for identifying new antibodies for novel targets. Next
39 generation sequencing (NGS) has recently emerged as a new method for the high throughput
40 characterization of IG and T cell receptor (TR) immune repertoires both *in vivo* and *in vitro*.
41 However, challenges remain for the NGS sequencing of scFv from combinatorial libraries owing to
42 the scFv length (> 800 bp) and the presence of two variable domains (VH and VL for IG) associated
43 by a peptide linker in a single chain. Here we show that single-molecule real-time (SMRT)
44 sequencing with the Pacific Biosciences (PacBio) RS II platform allows for the generation of full-
45 length scFv reads obtained from an *in vivo* selection of scFv-phages in an animal model of
46 atherosclerosis. We first amplified the DNA of the phagemid inserts from scFv-phages eluted from
47 an aortic section at the third round of the *in vivo* selection. From this amplified DNA, 450,558 reads
48 were obtained from 15 SMRT Cells. Highly accurate circular consensus sequences (CCS) from these
49 reads were generated, filtered by quality and then analysed by IMGT/HighV-QUEST with the
50 functionality for scFv. Full-length scFv were identified and characterized in 348,659 reads. Full-
51 length scFv sequencing is an absolute requirement for analysing the associated VH and VL domains
52 enriched during the *in vivo* panning rounds. In order to further validate the ability of SMRT
53 sequencing to provide high quality, full-length scFv sequences, we tracked the reads of an scFv-
54 phage clone P3 previously identified by biological assays and Sanger sequencing. Sixty P3 reads
55 showed 100% identity with the full-length scFv of 767 bp, 53 of them covering the whole insert of
56 977 bp, which encompassed the primer sequences. The remaining seven reads were identical over a
57 shortened length of 939 bp that excludes the vicinity of primers at both ends. Interestingly these reads
58 were obtained from each of the 15 SMRT Cells. Thus, the SMRT sequencing method and the
59 IMGT/HighV-QUEST functionality for scFv provides a straightforward protocol for characterization
60 of full-length scFv from combinatorial phage libraries.

61

62 **1 Introduction**

63 Immunoglobulin (IG) or antibody fragments displayed as single chain Fragment variable (scFv) on
64 filamentous phages (scFv-phages) are classically selected from scFv-phage combinatorial libraries to
65 obtain human antibodies specific for a given target (1-3). This selection from scFv-phage display
66 libraries is widely used for the discovery of novel specificities for therapeutic antibodies in cancer,
67 cardiovascular, autoimmune, infectious or neurodegenerative pathologies, with many of them at
68 various stages of clinical or research development (4-10). Classical *in vitro* phage display approaches
69 involve multiple rounds of selection (or panning) for the enrichment of scFv-phages that demonstrate
70 the desired specificity against a target followed, at the last selection round, by functional screening
71 and characterization of selected candidates using appropriate assays. At this very last step, analysis of
72 the selected scFv via Sanger sequencing is commonly used to identify sequences of interest, taking
73 advantage of the genotype-phenotype linkage inherent to the display system. A critical limitation of
74 using biological assays followed by Sanger sequencing is that only a minute fraction of the selected
75 library is actually sampled, a few hundred at best, whereas the selected library may usually contain
76 up to 10^5 to 10^6 variants. This limitation is further enhanced when scFv-phage selection is performed
77 *in vivo* (biopanning) in different pathological models in which scFv-phages can encounter a very
78 large panel of unknown biomarkers (11-13). Currently available next generation sequencing (NGS)

79 platforms allow the simultaneous sequencing of millions of reads. However, a main challenge for the
80 NGS sequencing of scFv from combinatorial libraries remains the scFv length > 800 bp, which is too
81 long for most NGS platforms. Up to now, NGS methods have only provided reads encompassing one
82 variable (V) domain (400 bp), therefore losing a critical piece of information found in scFv
83 sequences, that of the association of two specific V domains (variable heavy (VH) and variable light
84 (VL) for the IG) by the peptide linker. Although a few approaches have been proposed, retrieving
85 information regarding V domain association has still not been solved (14-16).

86 The analysis of antibody scFv sequences is a difficult exercise because not only are scFv composed
87 of two V domains, but these two V domains are different from each other and each can potentially be
88 extremely diverse. Indeed, the huge diversity of IG or antibodies results from complex *in vivo*
89 mechanisms that occur during the synthesis of the VH and VL domains, which include the molecular
90 rearrangements at the DNA level of the variable (V), diversity (D) (only for VH) and joining (J)
91 genes with nucleotide deletions and insertions (N-diversity) at the V-(D)-J junctions in the bone
92 marrow pre-B and immature B cells (17, 18). In spleen and lymph nodes, somatic hypermutations
93 accumulate in the mature B cell VH and VL, creating a huge diversity of the B cell membrane IG for
94 the recognition of foreign antigens. Following a specific antibody-antigen interaction the B cell
95 proliferates and generates clones engaged in *in vivo* selection and affinity maturation. The specificity
96 of the V domains is conferred by the complementarity determining regions (CDR) and more
97 particularly the CDR3 (19-21). The same features are observed in *in vitro* combinatorial libraries,
98 which mimic the natural *in vivo* diversity, selection and affinity maturation (1-3).

99 In order to manage, analyse and compare the extraordinary diversity of the immune repertoires,
100 IMGT®, the international ImMunoGeneTics information system®¹ (22), was created in 1989 in
101 Montpellier by Marie-Paule Lefranc (Montpellier University, CNRS), which marked the birth of
102 immunoinformatics (18), a new science at the interface between immunogenetics and bioinformatics.
103 IMGT® has developed online tools that provide a detailed and accurate analysis of the V domains
104 and which, in the case of nucleotide sequences, include IMGT/V-QUEST (23-25) for the analysis of
105 the rearranged V-(D)-J sequences of the IG or antibodies and T cell receptors (TR), and
106 IMGT/JunctionAnalysis (26, 27) for the analysis of the V-(D)-J junctions and of the included CDR3.
107 The algorithms and IMGT reference directories of these tools have been implemented in
108 IMGT/HighV-QUEST (28-31), the first and only web portal for NGS sequence analysis of IG and
109 TR, begun in 2010. IMGT/HighV-QUEST analyses up to 500,000 NGS reads per batch and includes
110 a statistical module for IMGT clonotype identification and comparison (analyses are performed on
111 the results of up to one million reads, from one or several batches) (30). IMGT/StatClonotype (32,
112 33), a stand-alone tool and R package, allows for the comparison of IMGT clonotype diversity and
113 expression between two NGS data sets, using the IMGT/HighV-QUEST statistical results output. In
114 order to overcome the analysis challenge of the NGS scFv, the IMGT/V-QUEST functionality
115 ‘Analysis of single chain Fragment variable (scFv)’ which includes the search and characterization of
116 two V domains in a single sequence (IMGT/V-QUEST Documentation² (34)) has recently been
117 integrated in IMGT/HighV-QUEST (IMGT/HighV-QUEST Documentation³).

118 The main challenge addressed in this study was to obtain high quality NGS reads of the scFv long
119 enough to encompass the two VH and VL domains and to analyse and characterize the association of
120 the two V domains in the NGS scFv reads using the scFv functionality in IMGT/HighV-QUEST.
121 Full-length scFv reads are expected to have a length of around 800 bp with a whole insert of around
122 1000 bp (including the 5’ and 3’ sequences of the vector and primer sequences at both ends). We
123 used the PacBio third-generation NGS technology, which provides long sequencing reads and the
124 highest consensus accuracy available today (35-38). For this project, the high accuracy is a result of
125 the generation of circular consensus sequencing (CCS) reads, by which the long sequencing reads

126 allow for multiple passes of the same insert and thus removal of random sequencing errors upon
127 consensus construction. Practically, one single molecule, real-time (SMRT) Cell has 150,000 zero-
128 mode waveguides (ZMW), of which 50,000-75,000 are loaded with single molecules during
129 sequencing, resulting in the production of ~50,000 to 75,000 unique consensus reads per run and per
130 SMRT Cell. PacBio NGS sequencing was performed from amplified DNA of scFv-phages isolated
131 from the third round of an *in vivo* biopanning in an animal model of atherosclerosis (12, 13). The
132 PacBio scFv CCS (version 2) reads were first analysed using IMG/HighV-QUEST with the scFv
133 functionality. In a second step, the sequence quality was evaluated by tracking the NGS reads of a
134 scFv-phage clone P3, identified by biological assays and Sanger sequencing.

135

136 **2 Materials and methods**

137 **2.1. *In vivo* selection of scFv-phages specific to atheroma plaque**

138 A fully human recombinant scFv antibody library (scFv cloned in the pMG72 phagemid vector,
139 containing the ampicillin resistant gene for the selection and maintenance of the phagemid) with a
140 diversity of 3.5×10^9 clones (full description in patent WO2007137616) was expressed by phage
141 display and selected *in vivo*, as previously described (13). The scFv-phages were obtained from the
142 scFv-phagemid combinatorial library by expression of the scFv on the phage surface, following the
143 addition of a helper filamentous phage to recombinant phagemid infected bacteria in the exponential
144 phase (39). Three rounds of biopanning were performed in atheromatous injured rabbits (12). All
145 animal experiments were performed in conformity with the Guide for the Care and Use of Laboratory
146 Animals (NIH Publication No. 85-23, revised 1996) and were accredited by the local ethical
147 committee (Animal Care and Use Committee of Bordeaux, France under the N°50120192). Briefly,
148 the procedure was the following (Figure 1): 2.4×10^{12} colony-forming units (cfu) of scFv-phages
149 were injected into an atheromatous rabbit. After 1h in circulation, the animal was sacrificed, the aorta
150 was retrieved and scFv-phages binding to the aorta were eluted in different fractions. The eluted
151 scFv-phages were re-amplified in XL1-Blue bacteria and following scFv expression at the phage
152 surface as above (39), the amplified scFv-phages were re-injected in another atheromatous animal.
153 Rounds 2 and 3 were conducted following the same procedure. The number of re-injected colony-
154 forming units were 4.8×10^{11} in round 2 and 3.9×10^{11} in round 3. The total number of eluted scFv-
155 phages from the third round was 1.5×10^7 cfu (total for seven fractions corresponding to different
156 areas of the aorta).

157 In this study, the analysed fraction is the one recovered after the third round of selection from the
158 endothelial cells of the damaged abdominal aorta vessel wall (named AAR3 for abdominal aorta
159 round 3) (Figure 2). The scFv-phages were amplified in XL1-Blue bacteria and plated on 145 mm
160 Petri dishes for storage of the whole AAR3 fraction before sequencing, and on 80 mm Petri dishes
161 for limiting dilution (quantification of recombinant bacteria and picking of individual clones) (Figure
162 2). The recombinant bacteria plated on 145 mm Petri dishes were scratched and stored at -80°C in
163 50% v/v glycerol. Around 3.5×10^5 clones issuing from the whole AAR3 fraction were counted by
164 limiting dilution of recombinant bacteria on 80 mm Petri dishes. Ninety-six recombinant bacteria
165 clones were individually picked and grown in selective medium on a 96-well MASTERBLOCK®
166 polypropylene storage plate (Greiner Bio-One, France) and stored for *in vitro* bioreactivity assays (*in*
167 *vitro* screening of scFv-phages on atheromatous proteins) and for Sanger sequencing.

168 Each of the 96 individual recombinant bacteria clones from the MASTERBLOCK® was re-picked
169 from an agarose 96-Well plate (Beckman Coulter, France). After incubation at 37°C overnight,
170 plasmid DNA extraction, PCR amplification and sequencing were then performed by Beckman
171 Coulter (France). One of these sequenced clones, P3 (767 bp), was used in this study for tracking the
172 PacBio reads that were identical or related to it.

173 **2.2. PacBio RS II sequencing of the whole AAR3 fraction**

174 **2.2.1. Generating high quality PCR products**

175 SMRT sequencing requires high-quality, doubled stranded DNA as input. To ensure this, plagemid
176 DNA of recombinant bacteria from the whole AAR3 fraction was extracted directly from the frozen
177 extract just before PCR amplification, using the QIAprep spin miniprep kit (Qiagen, France)
178 according to manufacturer's instructions. To generate clean, undamaged and non-chimeric
179 amplicons, the highest fidelity polymerase was used. All PCR reactions were performed in volumes
180 of 50 µL using 25 µL of the KAPA HiFi™ HotStart ready Mix (Kapa Biosystems, France) and 20 ng
181 of DNA template. Each primer was used at a final concentration of 0.3 µM. PCR reactions were
182 performed with the forward primer (Primer 1, 23-mer FWD) 5'-
183 TGCAAATTCTATTTCAAGGAGAC-3' and the reverse primer (Primer 2, 20-mer REV) 5'-
184 TCACGTGCAAAGCAGCGGC-3'. These primers were designed based on the plagemid vector;
185 for primer 1 from position -96 to -74 upstream of the BamHI site and for primer 2 from position 95 to
186 114 downstream of the NotI site (Figure 3A).

187 The PCR cycling protocol was chosen according to the manufacturer's instructions and consisted of
188 95°C for 3 min; 15 cycles of 98°C 20 s, 65°C 15 s, 72°C 30 s; one 72°C 1 min followed by one cycle
189 at 4°C. The reduction of the number of PCR cycles minimizes the PCR bias, however when further
190 decreases in the number of PCR cycles were attempted, it led to several contaminating, off-target
191 bands on an agarose gel.

192 The required quantity in PacBio guidelines for insert sizes superior to 750 bp is 500 ng of DNA. Four
193 PCR reactions were performed on the same AAR3 fraction for reproducibility purposes. Agarose gel
194 electrophoresis was used to confirm amplification, correct fragment size, and to check for non-
195 specific product contamination. A sizing marker was included to confirm size specificity (GeneRuler
196 1 kb Plus DNA Ladder, Thermo Fisher Scientific, France). Amplicons were cleaned using 1x ratio of
197 AMPure PB Beads (Pacific Biosciences). DNA purity and quantification (sample volume, yields and
198 size distributions) were evaluated and measured using the Agilent 2100 Bioanalyzer DNA12000 kit
199 (Agilent technologies).

200 **2.2.2. SMRTbell template preparation**

201 SMRTbell templates (PacBio, CA, United States) were constructed following the standard Amplicon
202 Sequencing Protocol⁴. The full procedure is explained in figure S1 (Supplementary data).

203 Briefly, PCR products were treated to repair DNA damage and then hairpin adapters were added via
204 blunt end ligation to produce SMRTbell templates using the SMRTbell template prep kit 1.0.
205 Exonucleases III and VII were used to remove failed ligation products and SMRTbell templates were
206 purified with AMPure PB Beads. The ratio of sequencing primer and polymerase was determined by
207 a PacBio calculator to correlate with SMRTbell concentrations and the 1,100-bp insert size. The
208 sequencing primer was annealed to the single-stranded loop region of the SMRTbell template, and
209 primer-annealed templates were then bound to DNA polymerase P6 using the DNA/polymerase

210 binding kit P6v2. The DNA-polymerase complexes were loaded on 15 SMRT Cells using MagBeads
211 onto the Pacbio RS II system and sequenced using the C4 chemistry and 6-hour movies.

212 Initial loading titrations were performed to identify the optimal loading concentration, identifying
213 0.03-0.035 nM as the best loading conditions for scFv PCR products. Each SMRT Cell generates
214 ~50,000 reads on average (considering that 50,000 to 75,000 ZMW can be optimally loaded with a
215 single molecule of DNA). A total of 15 SMRT Cells were loaded for the four different PCR products
216 of the same AAR3 fraction (three SMRT for sample 1 and four SMRT for samples 2, 3 and 4) to
217 cover, per PCR sample, the diversity of the AAR3 fraction (Figure S2, Supplementary data). The use of
218 15 SMRT cells was chosen to provide a thorough and sound proof-of-concept and read comparison
219 from four different PCR amplicons generated from the same fraction (AAR3).

220 **2.2.3. PacBio Circular Consensus Sequencing (CCS) read generation**

221 The PacBio RSII instrument produces sequencing reads with an average read length of ~15 kb, which
222 would theoretically pass over a 1,100 bp molecule more than 10 times, producing CCS2 sequences
223 with an accuracy ~99.9%. Longer read lengths can be achieved by increasing the instrument run
224 time, so these data were collected using 15 SMRT Cells and 6 hour movies, to maximize read length
225 and, thus, number of passes. Following sequencing, the raw data were processed using the CCS2
226 pipeline (CCS code used available at <https://github.com/PacificBiosciences/unanimity>). All CCS2
227 reads that were 99.9% accurate or greater were exported for further analysis. The raw NGS data of
228 the 85-related P3 sequences can be found in the NCBI Sequence Read Archive with the accession
229 number SRP124616.

230 **2.3. IMG/HighV-QUEST analysis of the PacBio CCS2 reads**

231 **2.3.1. Characterizing IMG/HighV-QUEST scFv reads from AAR3**

232 The FASTQ files of the PacBio CCS2 reads were imported and converted to FASTA sequences for
233 submission to IMG/HighV-QUEST (<http://www.imgt.org/HighV-QUEST/login.action>) (28-31),
234 which implements IMG/V-QUEST program version 3.4.2 (4 August 2016) and IMG/V-QUEST
235 reference directory release 201631-4. The analysis was performed with the advanced functionality
236 “Analysis of single chain Fragment variable (scFv)” (IMG/HighV-QUEST Documentation³).

237 Data filtering was applied with the following criteria to be fulfilled for each of the two V domains: (i)
238 >85% of identity of the V-REGION of the V domain with the V-REGION of the closest germline
239 IMG gene and allele, and (ii) in-frame V-(D)-J junction. Filtered sequences were then analysed to
240 identify the closest V, D (for VH) and J IMG genes and alleles, to characterize the amino acid (AA)
241 junction, to evaluate the mutations and to give a complete description of the scFv with IMG labels
242 (IMG Index > Single chain Fragment variable (scFv)⁵).

243 **2.3.2. Tracking and analysis of identical and related PacBio reads of the AAR3 scFv-phage 244 clone P3**

245 In order to evaluate the PacBio scFv read sequencing quality, reads identical or closely related to the
246 sequence of a scFv-phage clone P3 (previously isolated from the same fraction AAR3 and Sanger
247 sequenced) were tracked among the total reads generated in the 15 SMRT Cells by two approaches.
248 First, reads potentially related to P3 were searched for based on the expected VH-VL characteristics
249 determined by IMG/HighV-QUEST (same V and J genes and alleles, and same AA junctions).

250 Secondly, a Blast search was performed to check if P3-related reads could have been missed by the
251 IMGT/HighV-QUEST filtering. The fasta headers of the sequences were modified to include the
252 sequence set identifier and a blast database was built (formatdb 2.2.26) from the accumulated reads
253 from the 15 SMRT Cells. The database search was performed using the blastn program (2.2.26), with
254 the combined following criteria: longest alignments on the P3 Sanger (767 bp) sequence, highest
255 identity percentage, and maximum number of 20 mismatches or indels. Among the extracted reads
256 only those fulfilling the P3 characteristics (in term of V and J genes, allele names and AA junctions)
257 were retained.

258

259 **3 Results**

260 **3.1. Generation and analysis of PCR products from the whole AAR3 fraction**

261 The unbiased characterization of scFv from phage-display combinatorial libraries, in conjunction
262 with sequencing on the PacBio RS II system, requires high-quality PCR products with undamaged,
263 clean and non-chimeric amplicons. Creating PCR protocols to generate products that are truly
264 representative of the starting cell population is a major challenge.

265 To achieve this aim, (1) the complete frozen AAR3 fraction was directly used, without amplification,
266 as the source of recombinant bacteria; (2) efforts were made to limit the number of amplification
267 cycles (the four samples subject to 15 cycles) in order to reduce quantitative distortions as well as
268 error rates due to PCR artifact (data not shown); and (3) different high-fidelity Taq polymerases were
269 tested so as to fit with our amplification system. The required quantity (500 ng) and the correct size
270 of amplicons (~1,000 bp) were obtained with the KAPA HiFi™ HotStart polymerase (Figure 3A).
271 The quality and quantity of amplicons was confirmed using an Agilent 2100 Bioanalyzer. After
272 confirmation of the purity (Figure 3B), PCR samples were sequenced on the PacBio platform.

273 **3.2. PacBio sequencing and CCS2 analysis**

274 Sequencing was done on the PacBio RS II system using P6-C4 sequencing chemistry and SMRTbell
275 libraries generated from DNA amplified from the AAR3 fraction (corresponding to 3.5×10^5 scFv-
276 phagemids). For each of the four PCR samples, 3 or 4 SMRT Cells were run to ensure adequate
277 sampling of the scFv as described in Materials and Methods (Figure S2, Supplementary materiel). A
278 total of 15 SMRT Cells were used to cover the diversity of the AAR3 fraction.

279 Circular consensus sequencing (CCS2) analysis of the 15 SMRT Cells produced 450,558 reads.
280 These reads were filtered to remove any double-loaded wells or other artifactual/lower accuracy
281 reads. The data obtained from filtered, post-CCS2 analysis represent the reads achieving 99.9%
282 accuracy or above, derived only from wells loaded with a single molecule. Any reads that did not
283 reach consensus coverage of QV30 (99.9% accurate) or above were filtered out. Those settings and
284 filters are built into the CCS2 pipeline that is available via the PacBio web-based analysis software
285 (SMRT Link). Thus, these 450,558 reads passed initial filtering with an average pass number of 24
286 and a quality score of minimum 99.94% accuracy (Figure 3C). Another contributing factor to data
287 quality and throughput was the use of the longest movie lengths possible at the time (6 hrs) to ensure
288 the longest read lengths for analysis.

289 **3.3. IMGT/HighV-QUEST analysis of the scFv PacBio reads**

290 The 450,558 FASTQ PacBio CCS2 reads were converted in FASTA sequences and analysed using
291 IMGT/HighV-QUEST, as described in Materials and Methods. After analysis, a total of 391,655
292 scFv ‘candidates’ (i.e. sequences with two V regions, IMGT label V-REGION) (86.93% of the
293 submitted PacBio reads) were identified (Table 1). The scFv sequences were then filtered according
294 to the criteria described in Materials and Methods (>85% of identity of the V-REGION of the V
295 domains with the V-REGION of the closest germline IMGT genes and alleles and in-frame V-(D)-J
296 junction, determined for both V domains). The threshold of 85% of identity is the standard filter for
297 classical IG repertoire analysis in IMGT/HighV-QUEST (28). Following this filtering, 348,659 full-
298 length scFv representing 89.02% of the filtered sequences were identified (Table 1). These scFv
299 reads include 346,934 VH-VL or VL-VH expected scFv sequences (Table S1, Supplementary
300 material). The other 1,725 scFv reads comprise 171 VH-VH (5 to 22 found per SMRT Cell) and
301 1,554 VL-VL (68 to 158 found per SMRT Cell). These combinations most probably occurred during
302 the construction of the original scFv-phage combinatorial library. Thus these results provide a useful
303 and detailed overview of the content of the scFv combinatorial library.

304 The IMGT/HighV-QUEST analysis of the scFv reads included identification of the closest V, D (if
305 VH) and J IMGT genes and alleles, characterization of the junction, evaluation of the mutations and
306 complete description of both V domains with IMGT labels^{2,5}. These results demonstrate that both
307 domains of the scFv reads sequenced by PacBio could be fully characterized with the functionality
308 for scFv.

309 **3.4. Tracking and analysis of PacBio reads identical or related to the P3 clone across 15 SMRT** 310 **Cells**

311 ScFv-phages issuing from the *in vivo* AAR3 selected fraction have been screened by a high-
312 throughput flow cytometry method against atherosclerotic rabbit proteins. Some of the selected
313 clones were then processed into scFv-Fc format in HEK293 cells and validated by
314 immunohistochemistry on atheromatous aorta sections of two animal models of atherosclerosis
315 (ApoE -/- mouse and New Zealand White rabbit submitted to hypercholesterolemic diet) and on
316 human endarterectomy biopsies. A scFv-phage clone, named P3, selected for its high accuracy in
317 biological assays in the two animal models and human biopsies was Sanger sequenced and patented
318 (EP 17306337.1). The whole scFv P3 sequence (Figure S3) was then tracked among the 346,934 VH-
319 VL and VL-VH reads (Table S1) from the 15 SMRT Cells, by the two approaches described in
320 Materials and Methods, in order to evaluate the PacBio scFv read sequencing quality.

321

322

323 **3.4.1. Analysis of PacBio reads identical to the P3 clone**

324 Sixty PacBio reads with 100% identity to the region aligned with the P3 Sanger sequence (767 bp)
325 were obtained by both approaches. Interestingly, the P3 scFv was identified in the top 100 of the
326 most abundant VH-VL associations in all the 15 data sets. These PacBio reads have a length of 977
327 bp, including the 5’ and 3’ regions and the primers. Fifty-three of the 60 reads have a 100% identity
328 on their full length of 977 bp. As these PacBio reads were obtained from the 15 SMRT Cells (Table
329 2), we can confidently say that no sequencing error was observed in the 51,781 bp of these 53 reads.

330 Seven reads showed 100% identity on a 939 bp length (Table 2). Interestingly their mutations are all
331 localized at the ends of the primers and/or in the immediate vicinity of the 3’ primer (Table 3). A

332 most probable explanation is that they occur during the sequencing polymerisation (the priming step
333 could be excluded as different mutations were observed in different SMRT Cells and no degenerate
334 bases were used in the primers). Ignoring the mutations in or next to the primers, no sequencing error
335 was observed in 6,573 nucleotides of the 7 reads (100% identity on 939 bp). Combining these results
336 with those of the 53 reads (100% identity on the full length of 977 bp), no sequencing error was
337 detected on 58,354 nucleotides for the 60 scFv reads.

338 **3.4.2. Analysis of PacBio CCS reads related to the P3 clone**

339 In addition to the 60 PacBio reads with a 100% identity (53 reads on 977 bp and seven reads on 939
340 bp) (Table 2), 25 ‘related’ P3 reads were identified (Figure S4, Supplementary data). These 25 P3-
341 related reads showed 15 different mutation types in the insert and, for each type, a limited number of
342 reads (1 to 6) (Table 4). All together, the 15 mutation types are described by 28 different individual
343 mutations. This heterogeneity is in sharp contrast with the 60 reads with a 100% identical insert.

344 The observation of related reads was expected as the scFv are from a phagemid combinatorial library
345 in which point mutations were initially introduced to mimic the IG somatic hypermutations. The
346 analysis of the 25 reads was therefore performed in an attempt to distinguish among them related
347 clones with potential biological interest (as reflecting the diversity of the original library) from reads
348 with artifactual differences.

349 For 15 of these PacBio CCS reads, the origin of the mutations (mostly substitutions) was clearly from
350 the scFv phagemid combinatorial library (and therefore of potential interest given their relatedness to
351 P3). This was supported by the fact that identical mutated reads were obtained from different PCR
352 and from different SMRT Cells (category A in Table 4). These included 7 reads (highlighted in pink)
353 with four identical individual mutations, 2 reads (highlighted in green) with three identical individual
354 mutations different from those described above and one shared by both groups (c838>g (VL)), 6
355 reads (highlighted in blue) with two identical individual mutations (and still different from those of
356 the previous mutation types). All the substitution mutations are localized in the VL, in agreement
357 with mutations of the VL and VH domains being targeted differently before the assembly into the
358 scFv and confirming that the differences observed are intrinsic to the VL domain, and not due to PCR
359 or sequencing errors. The IMGT/V-QUEST alignment of the 15 PacBio NGS sequences with the
360 initial P3 Sanger sequence is provided as an additional PDF file in supplementary data (Figure S5,
361 Supplementary data).

362 For the other 10 PacBio reads, present in one copy (category B in Table 4), the origin of the
363 differences could not be determined with certainty: and any of the possible explanations: putative
364 sequencing errors, PCR amplification errors or diversity of the combinatorial library could not be
365 formally excluded at this stage. The 1nt-deletion observed in VH and VL for reads 85 and 40
366 respectively, the large nt deletion at the 3’end in read 78, the nt deletion at both 5’ and 3’ ends for
367 read 81 and the one or two nt insertions in VH or VL for reads 84 and 9 could be considered as
368 sequencing errors. The six reads with different single substitution and the one with 3 mutations could
369 be attributed to PCR amplification errors or diversity of the combinatorial library. Moreover, and
370 even if these differences are due to diversity of the combinatorial library, their single copy number
371 suggests that they are from unselected (or poorly selected) scFv-phages from the library background
372 and could be ignored in the library screening.

373

374

375 4. Discussion

376 Antibody libraries are important resources to derive antibodies to be used for a wide range of
377 diagnostic and therapeutic applications. In vivo or ex vivo phage-display selections have emerged as
378 interesting ways to identify accurate antibodies in the context of the pathologic microenvironment
379 (41). Although advancements in automation of biological assays have greatly improved screening
380 strategies, high-throughput campaigns are still severely limited in the number of antibody fragments
381 that can be interrogated, providing only a tiny fraction of the enriched phage library. Access to the
382 genetic information encoded in antibody repertoires by next-generation sequencing (NGS) should
383 allow more in depth analysis of the diversity of the selected library. However, the currently available
384 NGS platforms that were capable of providing several million reads per run, generated only short
385 reads of up to 300-700 bp (42). Therefore, only synthetic combinatorial scFv libraries, in which
386 diversity was confined to CDR3 regions of the heavy and light chains, have benefited from NGS
387 potential in the extensive in silico analysis of complex collections of selected antibodies (43).

388 While Mi-Seq, 454 or Ion-Torrent technologies will completely sequence heavy and light variable
389 domains, they are currently insufficient to cover the full-length scFv, which are comprised of both
390 VH and VL domains, connected by a peptide linker. These technologies often necessitate consensus
391 building of multiple reads originating from the scFv fragment to obtain whole sequence information
392 (44-46). Therefore, NGS sequencing of scFv fragments longer than 800 bp is still hampered by
393 technical limitations in the length of reads. These limitations could be circumvented by the third
394 generation PacBio sequencing platform. Capitalizing on PacBio SMRT DNA sequencing for high-
395 resolution and high-throughput HLA typing (36,37), we develop here a PacBio SMRT/CCS2
396 approach, combined with IMGT/HighV-QUEST analysis of full-length scFv reads to provide a
397 straightforward protocol for characterization of the complete VH and VL domains of scFv from fully
398 human combinatorial libraries.

399 PacBio SMRT sequencing generated 450,558 reads of about 1000 bp across 15 SMRT Cells,
400 following DNA amplification of the scFv insert from 3.5×10^5 in vivo selected scFv-phages.
401 IMGT/HighV-QUEST with its scFv functionality allowed for further filtration and characterization
402 of the 348,659 PacBio CCS reads as containing full-length scFv, which represent 89.02% of the
403 overall filtered sequences from the 15 SMRT Cells run. Among them, 346,934 identified expected
404 VH-VL or VL-VH full-length scFv reads.

405 In order to evaluate the PacBio scFv read sequencing quality, a selected scFv-phage clone P3,
406 previously identified from the AAR3 fraction by biological screenings and sequenced by the Sanger
407 methodology, was tracked within the 346,934 VH-VL and VL-VH scFv reads characterized by
408 IMGT/HighV-QUEST. Sixty PacBio CCS2 reads were identified from the 15 SMRT Cells that
409 demonstrated 100% identity on a length of 939 bp (including the complete scFv of 767 bp and the 5'
410 and 3' regions), and for 53 of them on the full length of 977 bp (including the primers at both ends).
411 Thus no sequencing error was observed on a total of 51,728 bp for these 53 reads, obtained from the
412 15 SMRT Cells (and 58,354 bp if the seven reads with 100% on 939 bp are included), validating the
413 capacity of the PacBio SMRT-CCS2 method to produce reads with an accuracy of > 99.9% for
414 complete sequences of inserts approaching 1000 bp. It is of interest to note that these PacBio reads
415 were obtained from the 15 SMRT cells.

416 In addition to the 60 PacBio reads with a 100% identity, 25 'related' P3 reads were identified and for
417 15 of them, mutation analysis (category A in Table 4) clearly showed that they are related to P3. This

418 was supported by the fact that identical mutated reads in the VL (in agreement with the diversity
419 generated during the library construction) were obtained from different PCR samples and from
420 different SMRT Cells. P3 (60 reads) with its 15 related reads are among the top 100 most represented
421 associated VH-VL domains in all the 15 data sets.

422 This study provides the first proof-of-concept that similar sequences could be tracked in phage-
423 display selected scFv samples and their frequency determined by the number of reads. The favourite
424 candidates chosen for their high frequency of enrichment could be rescued from these in silico data
425 for implementation of downstream biological assays. This could be easily done by custom gene
426 chemical synthesis, which offers the utmost flexibility and efficiency with high production yields
427 (47).

428

429 Thus the large number of sequenced reads delivered, following the enrichment process, could be
430 ideally suited for a more extensive evaluation of antibody candidates by biological assays. In that
431 way, bringing full sequence data from NGS will accelerate search for identification of both the
432 antibodies and their targeted biomarkers. We thus aimed to combine the sensitivity of the sequencing
433 approach with the functional information provided by the immune assays. This in silico approach
434 could be applied to any other pathology and phage-display screening methodology. Resolving the
435 issue of complete scFv sequencing has the potential to profoundly impact the selection process of
436 antibodies with desired properties after phage display biopanning with a special focus of in vivo
437 selections. It is expected that this will contribute to therapeutic antibody discovery, selection and
438 development.

439 Limitations:

440 Scientists working on immunoglobulins are very concerned by the problems of sequencing errors
441 versus mutations and of short read assembly. This study demonstrates that the PacBio SMRT-CCS2
442 method is able to produce reads with an accuracy of > 99.9% for complete sequences of scFv inserts
443 without the need of in silico VH and VL assembly as usually necessary using MiSeq or Ion Torrent
444 technologies (45,46). To assess the read quality of PacBio sequencing, we performed a pilot study
445 using different PCR samples and 3 or 4 SMRT for each PCR, starting from the same AAR3 fraction.
446 Using the biologically validated P3 clone as a reference, we demonstrated that P3 clone is among the
447 top 100 in vivo selected clones with a representativity of 0.025%. Moreover, among the 85 P3 related
448 sequences, 60 were 100% identical and 15 clearly originated from the scFv phagemid combinatorial
449 library. From the IMGT/V-QUEST alignment of P3 pink, blue and green groups of reads with the
450 initial P3 Sanger sequence, we can confidently say that 88% of reads were free of sequencing errors.
451 However, a doubt remains for 10 of them, present in just one copy. The nt-deletions or nt insertions
452 observed in VH and VL for six of the 10 reads could be definitively considered as sequencing errors.
453 For the single substitutions observed in seven reads, it is obviously impossible to determine their
454 origin. To circumvent this limitation, other technologies could be considered, such as inverse PCR
455 method based on VH CDR3 or VL CDR3 sequences (46) to synthesize these clones from the AAR3
456 fraction. Nonetheless, what is of utmost importance in our study is to identify over-represented
457 clones (from the top 100 candidates) and to proceed to the rescue of highly-enriched scFv and not
458 isolated clones. Indeed, during phage-display selections, the reads of greatest interest will have the
459 greatest depth of coverage, having expanded in the pool and thus receiving greater proportional read
460 depth.

461

462

463

464 **Conflict of Interest Statement**

465 C.H., N.L.H, J.K. and J.K. are full-time employees at Pacific Biosciences, a company developing
466 single-molecule sequencing technologies. M.S. was a full-time employee at Pacific Biosciences at
467 the time of this work. P.M. is full-time employee at LFB, a pharmaceutical group specialising in
468 biological medicinal products. The other authors declare that they have no conflict of interest.
469

470 **Author Contributions**

471 AH, FO, MJJV, JLT, MS, SK, MPL and GCS conceived and designed the experiments. PM provided
472 the library. MS sequenced the data. NLH processed the raw data. VG designed the analysis
473 algorithms and PD implemented the tool for NGS. CH, JK, AG and JK supervised the sequencing
474 procedure. GCS and MPL supervised the project. AH, MS, GCS and MPL wrote the paper. All the
475 authors have read and approved the final manuscript.
476

477 **Funding**

478 IMGT® is currently supported by the Centre National de la Recherche Scientifique (CNRS); the
479 Ministère de l'Enseignement Supérieur et de la Recherche (MESR); the Montpellier University,
480 France; the Agence Nationale de la Recherche (ANR) Labex MabImprove [ANR-10-LABX-5301];
481 BioCampus Montpellier; Région Languedoc-Roussillon (Grand Plateau Technique pour la Recherche
482 (GPTR)). This work was granted access to the HPC@LR and to the High Performance Computing
483 (HPC) resources of the Centre Informatique National de l'Enseignement Supérieur (CINES) and to
484 Très Grand Centre de Calcul (TGCC) of the Commissariat à l'Energie Atomique et aux Energies
485 Alternatives (CEA) under the allocation [036029] (2010-2017) made by GENCI (Grand Equipement
486 National de Calcul Intensif). This study was funded by a public grant from the French National
487 Research Agency within the context of the Investments for the Future Program, referenced ANR-10-
488 LABX-0057 and named TRAIL. A public grant from ANR-13-BSV5-0018 SVSE5 program, named
489 ATHERANOS supported this work.

490

491 **Abbreviations**

492 CCS: circular consensus sequencing, IG: immunoglobulin, IMGT: IMGT®, the international
493 ImMunoGeneTics information system®, NGS: next generation sequencing, scFv: single chain
494 Fragment variable, SMRT: single-molecule real-time, TR: T cell receptor

495

496 **Acknowledgments**

497 We would like to thank Franck Salin for his helpful advice on DNA amplification and analysis using
498 the Agilent 2100 Bioanalyzer. We are really grateful to Alexandre Fontayne for his help in the P3
499 patenting process. We thank Roberto Lleras for helpful information. We are grateful to Gérard
500 Lefranc for helpful comments, Arthur Lavoie and Karthik Kalyan for IMGT/HighV-QUEST and to
501 all IMGT team members for their constant motivation. IMGT® is Academic Institutional Member of
502 the International Medical Informatics Association (IMIA) and of the Global Alliance for the
503 Genomics and Health (GA4GH).

504

505 **Footnotes**506 1. <http://www.imgt.org>507 2. http://www.imgt.org/IMGT_vquest/share/textes/imgtvquest.html508 3. <http://www.imgt.org/HighV-QUEST/doc.action>509 4. <http://www.pacb.com/wp-content/uploads/Procedure-Checklist-Amplicon-Template-Preparation-and-Sequencing.pdf>510 5. <http://www.imgt.org/IMGTindex/scFv.php>511 6. <http://www.imgt.org/IMGTScientificChart/Nomenclature/IMGTmutation.html>512 **References**513 514 1. McCafferty J, Griffiths AD, Winter G, Chiswell DJ. Phage antibodies: filamentous phage
515 displaying antibody variable domains. *Nature* (1990) **348**(6301):552–4. doi: 10.10138/348552a0516 2. Marks JD, Hoogenboom HR, Bonnert TP, McCafferty J, Griffiths AD, Winter G. By-passing
517 immunization. Human antibodies from V-gene libraries displayed on phage. *J Mol Biol.* (1991)
518 **222**(3):581-97.519 3. Griffiths AD, Malmqvist M, Marks JD, Bye JM, Embleton MJ, McCafferty J, et al. Human anti-
520 self antibodies with high specificity from phage display libraries. *EMBO J.* (1993) **12**(2):725-34.521 4. Peneff C, Lefranc M-P, Dariavach P. Characterisation and specificity of two single-chain Fv
522 antibodies directed to the protein tyrosine kinase Syk. *J Immunol Methods* (2000) **236**:105-115.523 5. Pelat T, Hust M, Laffly E, Condemine F, Bottex C, Vidal D, et al. A high affinity, human-like
524 antibody fragment (scFv) neutralising the lethal factor (LF) of *Bacillus anthracis* by inhibiting PA-
525 LF complex formation. *Antimicrob Agents Chemother* (2007) **51**:2758-2764. doi:
526 10.1128/AAC.01528-06527 6. Thie H, Meyer T, Schirrmann T, Hust M, Dübel S. Phage display derived therapeutic antibodies.
528 *Curr Pharm Biotechnol* (2008) **9**:439–446.529 7. Marcus WD, Wang H, Lindsay SM, Sierks MR. Characterization of an antibody scFv that
530 recognizes fibrillar insulin and beta-amyloid using atomic force microscopy. *Nanomedicine* (2008)
531 **4**(1):1-7. doi: 10.1016/j.nano.2007.11.003532 8. Pelat T, Hust M, Hale M, Lefranc M-P, Dübel S, Thullier P. Isolation of a human-like antibody
533 fragment (scFv) that neutralizes ricin biological activity. *BMC Biotechnol* (2009) **9**:60.
534 doi:10.1186/1472-6750-9-60.

- 535 9. Tian H, Davidowitz E, Lopez P, He P, Schulz P, Moe J, et al. Isolation and characterization of
536 antibody fragments selective for toxic oligomeric tau. *Neurobiol Aging* (2015) **36**(3):1342-55. doi:
537 10.1016/j.neurobiolaging.2014.12.002
- 538 10. Williams SM, Venkataraman L, Tian H, Khan G, Harris BT, Sierks MR. Novel atomic force
539 microscopy based biopanning for isolation of morphology specific reagents against TDP-43 variants
540 in amyotrophic lateral sclerosis. *J Vis Exp* (2015) Feb 12;(96). doi: 10.3791/52584.
- 541 11. Krag DN, Shukla GS, Shen GP, Pero S, Ashikaga, T., Fuller, S., et al. Selection of tumor-binding
542 ligands in cancer patients with phage display libraries. *Cancer Res* (2006) **66**:7724-7733.
- 543 12. Deramchia K, Jacobin-Valat MJ, Laroche-Traineau J, Bonetto S, Sanchez S, Dos Santos P, et al.
544 By-passing large screening experiments using sequencing as a tool to identify scFv fragments
545 targeting atherosclerotic lesions in a novel *in vivo* phage display selection. *Int J Mol Sci* (2012)
546 **13**:6902-6923. doi: 10.3390/ijms13066902
- 547 13. Deramchia K., Jacobin-Valat, MJ, Vallet A, Bazin, H., Santarelli, X., Sanchez, S., et al. *In vivo*
548 phage display to identify new human antibody fragments homing to atherosclerotic endothelial and
549 subendothelial tissues. *Am J Pathol* (2012) **180**:2576-2589. doi: 10.1016/j.ajpath.2012.02.013.
550 Erratum in: *Am J Pathol*. 2012 Nov;181(5):1889.
- 551 14. Ravn U, Gueneau F, Baerlocher L, Osteras M, Desmurs M, Malinge P, et al.. By-passing in vitro
552 screening--next generation sequencing technologies applied to antibody display and *in silico*
553 candidate selection. *Nucleic Acids Res* (2010) **38**(21):e193. doi: 10.1093/nar/gkq789
- 554 15. Larman HB, Xu GJ, Pavlova NN, Elledge SJ. Construction of a rationally designed antibody
555 platform for sequencing-assisted selection. *Proc Natl Acad Sci U S A* (2012) **109**:18523-18528. doi:
556 10.1073/pnas.1215549109
- 557 16. Ravn U, Didelot G, Venet S, Ng KT, Gueneau, F, Rousseau F, et al. Deep sequencing of phage
558 display libraries to support antibody discovery. *Methods* (2013) **60**(1):99-110. doi:
559 10.1016/j.ymeth.2013.03.001
- 560 17. Lefranc M-P, Lefranc G. *The Immunoglobulin FactsBook*. London: Academic Press (2001). pp 1-
561 458.
- 562 18. Lefranc M-P. Immunoglobulin and T cell receptor genes: IMGT(®) and the birth and rise of
563 Immunoinformatics. *Front Immunol* (2014) **5**:22. doi: 10.3389/fimmu.2014.0002223.
- 564 19. Martinez O, Gangi E, Mordi D, Gupta S, Dorevitch S, Lefranc M-P, et al. Diversity in the
565 Complementarity Determining Region 3 (CDR3) of Antibodies from mice with evolving anti-TSHR
566 antibody responses. *Endocrinology* (2007) **148**:752-761. doi: 10.1210/en.2006-1096
- 567 20. Li L, Wang XH, Williams C, Volsky B, Steczko O, Seaman MS, et al. A broad range of
568 mutations in HIV-1 neutralizing human monoclonal antibodies specific for V2, V3, and the CD4
569 binding site. *Mol Immunol* (2015) **66**(2):364-374. doi: 10.1016/j.molimm.2015.04.011
- 570 21. Marillet S, Lefranc M-P, Boudinot P, Cazals F. Novel structural parameters of Ig-Ag complexes
571 yield a quantitative description of interaction specificity and binding affinity. *Front Immunol* (2017)
572 **8**:34. doi: 10.3389/fimmu.2017.00034
- 573 22. Lefranc M-P, Giudicelli V, Duroux P, Jabado-Michaloud J, Folch G, Aouinti S, et al. IMGT®,
574 the international ImMunoGeneTics information system® 25 years on. *Nucleic Acids Res* (2015)
575 **43**:D413-22. doi: 10.1093/nar/gku1056

- 576 23. Giudicelli V, Chaume D, Lefranc M-P. IMGT/V-QUEST, an integrated software program for
577 immunoglobulin and T cell receptor V-J and V-D-J rearrangement analysis. *Nucleic Acids Res* (2004)
578 **32**:W435–40. doi:10.1093/nar/gkh412
- 579 24. Brochet X, Lefranc M-P, Giudicelli V. IMGT/V-QUEST: the highly customized and integrated
580 system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res* (2008)
581 **36**:W503-8. doi:10.1093/nar/gkn316
- 582 25. Giudicelli V, Brochet X, Lefranc M-P. IMGT/V-QUEST: IMGT standardized analysis of the
583 immunoglobulin (IG) and T cell receptor (TR) nucleotide sequences. *Cold Spring Harb Protoc*
584 (2011) **6**:695–715. doi:10.1101/pdb.prot5633
- 585 26. Yousfi Monod M, Giudicelli V, Chaume D, Lefranc M-P. IMGT/JunctionAnalysis: the first tool
586 for the analysis of the immunoglobulin and T cell receptor complex V-J and V-D-J JUNCTIONs.
587 *Bioinformatics* (2004) **20**:i379-85. doi:10.1093/bioinformatics/bth945.
- 588 27. Giudicelli V, Lefranc M-P. IMGT/JunctionAnalysis: IMGT standardized analysis of the V-J and
589 V-D-J junctions of the rearranged immunoglobulins (IG) and T cell receptors (TR). *Cold Spring*
590 *Harb Protoc* (2011) **6**:716–25. doi:10.1101/pdb.prot5634
- 591 28. Alamyar E., Giudicelli V, Li S, Duroux P, Lefranc M-P. IMGT/HighV-QUEST: the IMGT® web
592 portal for immunoglobulin (IG) or antibody and T cell receptor (TR) analysis from NGS high
593 throughput and deep sequencing. *Immunome Res* (2012) **8**:1:2. doi:10.1007/978-1-61779-842-9_32
- 594 29. Alamyar E, Duroux P, Lefranc M-P, Giudicelli V. IMGT® tools for the nucleotide analysis of
595 immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG
596 mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. In: Christiansen F, Tait B, editors.
597 *Methods Mol Biol*. New York: Humana Press, Springer (2012) **882**:569–604. doi:10.1007/978-1-
598 61779-842-9_32
- 599 30. Li S, Lefranc M-P, Miles JJ, Alamyar E, Giudicelli V, Duroux P, et al. IMGT/HighV-QUEST
600 paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire
601 immunoprofiling. *Nat Commun* (2013) **4**:2333 doi: 10.1038/ncomms3333
- 602 31. Giudicelli V, Duroux P, Lavoie A, Aouinti S, Lefranc M-P, Kossida S. From IMGT-
603 ONTOLOGY to IMGT/HighV-QUEST for NGS immunoglobulin (IG) and T cell receptor (TR)
604 repertoires in autoimmune and infectious diseases. *Autoimmun Infec Dis* (2015) **1**:1. doi:
605 10.16966/aidoa.103
- 606 32. Aouinti S, Malouche D, Giudicelli V, Kossida S, Lefranc M-P. IMGT/HighV-QUEST Statistical
607 significance of IMGT clonotype (AA) diversity per gene for standardized comparisons of next
608 generation sequencing immunoprofiles of immunoglobulins and T cell receptors. *PLoS One* (2015)
609 **10**(11):e0142353. doi: 10.1371/journal.pone.0142353
- 610 33. Aouinti S, Giudicelli V, Duroux P, Malouche D, Kossida S, Lefranc M-P. IMGT/StatClonotype
611 for pairwise evaluation and visualization of NGS IG and TR IMGT clonotype (AA) diversity or
612 expression from IMGT/HighV-QUEST. *Front Immunol* (2016) **7**:339. doi:
613 10.3389/fimmu.2016.00339
- 614 34. Giudicelli V, Duroux P, Kossida S, Lefranc M-P. IG and TR single chain Fragment variable
615 (scFv) sequence analysis: a new advanced functionality of IMGT/V-QUEST and IMGT/HighV-
616 QUEST. *BMC Immunology* (2017) **18**(1):35. doi: 10.1186/s12865-017-0218-8.
- 617 35. Rhoads A, Au, KF. PacBio Sequencing and Its Applications. *Genomics Proteomics*
618 *Bioinformatics* (2015) **13**:278-289. doi: 10.1016/j.gpb.2015.08.002

- 619 36. Mayor NP, Robinson J, McWhinnie AJ, Ranade S, Eng K, Midwinter W, et al. HLA Typing for
620 the Next Generation. *PloS One* (2015) **10**, e0127153. doi: 10.1371/journal.pone.0127153.
- 621 37. Westbrook CJ, Karl JA, Wiseman RW, Mate S, Koroleva G, Garcia K, et al. No assembly
622 required: full-length MHC class I allele discovery by PacBio circular consensus sequencing. *Hum*
623 *Immunol* (2015) **76**:891-896. doi: 10.1016/j.humimm.2015.03.022
- 624 38. Carapito R., Radosavljevic M., Bahram S. Next-Generation Sequencing of the HLA locus:
625 Methods and impacts on HLA typing, population genetics and disease association studies. *Hum*
626 *Immunol* (2016) **77**:1016-1023. doi: 10.1016/j.humimm.2016.04.002
- 627 39. Hua TD, Souriau C, Marin M, Lefranc M-P, Weill M. Construction d'un répertoire de fragments
628 d'anticorps scFv et son expression à la surface de phages filamenteux [French]. In: Lefranc and
629 Lefranc (1997) *Ingénierie des anticorps. Banques combinatoires*. Les Editions INSERM, 1997, Paris.
630 Techniques en Immunologie. ISBN 2-85598-687-7. Chap 5, 39-54
- 631 40. Lefranc M-P. IMGT Locus on Focus: A new section of Experimental and Clinical
632 Immunogenetics, *Exp Clin Immunogenet* (1998) **15**:1-7.
- 633 41. Krag DN, Shukla GS, Shen GP, Pero S, Ashikaga T, Fuller S, et al. (2006) Selection of tumor-
634 binding ligands in cancer patients with phage display libraries. *Cancer Res* **66**: 7724-7733
- 635 42. Glanville J, D'Angelo S, Khan TA, Reddy ST, Naranjo L, Ferrara F, et al. (2015) Deep
636 sequencing in library selection projects: what insight does it bring? *Curr Opin Struct Biol* **33**: 146-
637 160
- 638 43. Ravn U, Didelot G, Venet S, Ng K, Gueneau F, Rousseau F, et al. (2013). Deep sequencing of
639 phage display libraries to support antibody discovery. *Methods*, **60**; 99-110
- 640 44. Yang W, Yoon A, Lee S, Kim S, Han J, Chung J (2017) Next-generation sequencing enables the
641 discovery of more diverse positive clones from a phage-displayed antibody library. *Exp Mol Med* **49**:
642 e308
- 643 45. DeKosky BJ, Kojima T, Rodin A, Charab W, Ippolito GC, Ellington AD, et al. (2015) In-depth
644 determination and analysis of the human paired heavy- and light-chain antibody repertoire. *Nat Med*
645 **21**: 86-91; 1
- 646 46. D'Angelo S, Kumar S, Naranjo L, Ferrara F, Kiss C, Bradbury AR (2014) From deep sequencing
647 to actual clones. *Protein Eng Des Sel* **27**: 301-307
- 648 47. Sai T Reddy, Xin Ge, Aleksandr E Miklos, Randall A Hughes, Seung Hyun Kang, Kam Hon Hoi
649 et al. (2010) Monoclonal antibodies isolated without screening by analyzing the variable-gene
650 repertoire of plasma cells. *Nat Biot* **28**: 965 - 971

651
652
653
654
655
656
657

658

659

660

661

662

663 Table 1. IMGT/HighV-QUEST analysis of the scFv PacBio reads.

664

PCR sample N°	Number of reads with 99.9% predicted accuracy	Mean number of passes	Number of movies	SMRT Cell N°	Number of PacBio CCS2 analysed reads	Number of scFv candidates in analysed reads	% of scFv candidates in analysed reads, i.e., coverage	Number of filtered-in reads	% of scFv in filtered-in reads, i.e., coverage	% of scFv in analysed reads, i.e., coverage
s1	91828	24	3	1	29,224	25,419	86.98	22,906	90.11	78.38
				2	32,240	28,120	87.22	25,228	89.72	78.25
				3	30,364	26,496	87.26	23,799	89.82	78.38
s2	129640	23	4	4	34,082	29,729	87.23	26,657	89.67	78.21
				5	33,510	29,213	87.18	26,407	90.39	78.80
				6	31,980	27,874	87.16	25,032	89.80	78.27
				7	30,068	26,289	87.43	23,695	90.13	78.80
s3	115446	24	4	8	34,890	30,183	86.51	26,990	89.42	77.36
				9	29,373	25,314	86.18	22,468	88.76	76.49
				10	26,465	22,776	86.06	20,044	88.00	75.74
				11	24,718	21,358	86.41	18,741	87.75	75.82
s4	113644	24	4	12	25,128	21,881	87.08	19,120	87.38	76.09
				13	23,693	20,515	86.59	17,756	86.55	74.94
				14	32,293	28,093	86.99	24,762	88.14	76.68
				15	32,530	28,395	87.29	25,054	88.23	77.02
Total	450,558				450,558	391,655	86.93	348,659	89.02	77.38

665

666 Each CCS read counts as “1X coverage” over the scFv molecules of interest. The coverage is given in percentage of
667 the scFv of interest (analysed reads, filtered-in reads and analysed reads).

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683
 684
 685
 686
 687
 688
 689
 690
 691
 692

Table 2. PacBio reads 100% identical to the aligned P3 Sanger sequence and 100% identical between them on 977 bp (53 reads) or 939 bp (7 reads).

PCR sample N°	Number of P3 PacBio reads per PCR sample	SMRT Cell N°	Number of P3 PacBio reads per SMRT Cell	100% on 977 bp (53 reads)	100% on 939 bp (7 reads)	GenBank/ENA/DDBJ accession number
s1	15	1	6	1, 2, 3, 4, 6, 8		MG272208
		2	5	10, 30	11, 12, 13	
		3	4	16, 17, 19	15	
s2	14	4	2	32, 33		
		5	2	36	35	
		6	3	37, 38, 39,		
		7	7	20, 41, 44, 45, 46, 47, 48		
s3	15	8	6	49, 50, 51, 52, 53, 54		
		9	4	56, 57, 59, 60		
		10	4	63, 65, 66	64	
		11	1	68		
s4	16	12	4	70, 72, 73	71	
		13	2	75, 76		
		14	3	21, 80, 83		
		15	7	23, 24, 25, 26, 27, 28, 29		
Total	60		60	53	7	

693
 694
 695
 696
 697
 698

699

700

701

702

703

704 Table 3. Mutations observed at the 5' and 3' end of the seven PacBio reads with 100 % identity on
 705 939bp (positions 3 to 941). Positions of the primers are 1-23 and 958-977.

PacBio read N° (assigned in the list 1-85)	PCR sample N°	SMRT Cell N°	Mutation description ^a	Mutation localization	GenBank/ENA/DDBJ accession number
13	s1	2	one 1nt-deletion (g2>del)	5' end of the 5' primer	MG272209
11	s1	2	two 1 nt-deletion (t975>del, a977>del)	3'end of the 3' primer	MG272210
15	s1	3	one 1 nt-substitution (c956>t)	vicinity of the 3' primer	MG272211
12	s1	2	one 1 nt-deletion (a942>del) ^b	vicinity of the 3'primer ^b	MG272212
35	s2	5			
71	s4	12			
64	s3	10	two 1 nt-deletion (a942>del) ^b , (t975>del)	vicinity of the 3'primer ^b , 3'end of the 3' primer	MG272213

706 ^a Mutations are described according to the IMGT Scientific chart rules⁶ (40).

707 ^b The 1 nt-deletion (a942>del) found in reads from the 4 samples most probably originates from the
 708 library.

709

710

711

712

713

714

715 **Table 4.** Mutation heterogeneity observed in 25 P3 related PacBio reads, in contrast with the 60 P3
716 identical PacBio reads with 100% identity on the complete scFv. Positions of the primers are 1-23
717 and 958-977. Category A: 15 P3-related reads of potential biological interest (mutations due to the
718 VL diversity originating from the combinatorial library). Pink, green and blue colours highlight
719 groups of reads with in common identical substitution mutations. Category B: 10 P3-related reads
720 with undefined origin of the mutations.

721

Read categories	PacBio read N° (assigned in the list 1-85)	PCR sample N°	SMRT Cell N°	Number of reads/mutation type	Mutation type	Mutation description ¹	GenBank/ENA/DDBJ accession number
A (15 reads)	7, 18, 43, 67, 69, 79	1, 2, 3, 4	1, 3, 7, 11, 12, 14	6	four 1 nt-substitution	a545>g (VL), g686>a (VL), a757>g (VL), c838>g (VL)	MG272218
	58	3	9	1	four 1 nt-substitution with, in 3', a large deletion	a545>g (VL), g686>a (VL), a757>g (VL), c838>g (VL), a886-a977>del (92nt)	MG272219
	61	3	9	2	four 1 nt-substitution	c741>t (VL), g837>a (VL), c838>g (VL), g843>t (VL)	MG272220
	74	4	13		four 1 nt-substitution with, at the 3' end of the 3' primer, a 1 nt-deletion	c741>t (VL), g837>a (VL), c838>g (VL), g843>t (VL), a977>del	MG272221
	5, 14, 31, 34, 42, 82	1, 2, 4	1, 2, 4, 7, 14	6	two 1 nt-substitution	c720>t (VL), t744>c (VL)	MG272223
B (10 reads)	85	2	4	2	one 1nt-deletion	c242>del (VH)	MG272216
	40	2	6		one 1nt-deletion	g600>del (VL)	MG272217
	77	4	13	6	one 1 nt-substitution	g495 >a (linker)	MG272227
	22	4	14		one 1 nt-substitution	t624>c (VL)	MG272224
	55	3	9		one 1 nt-substitution	a627>g (VL)	MG272225
	62	3	10		one 1 nt-substitution	g736>a (VL)	MG272226
	78	4	13		one 1 nt-substitution with, at the 3' end of the 3' primer, a 3 nt-deletion	t599>g (VL), t975-a977>del (3nt)	MG272228

81	4	14		one 1 nt-substitution with, at the 5' end of the 5' primer, a 1 nt-deletion, and at the 3' end, a 19-nt primer deletion	<i>g2>del, a715>g (VL), c959-a977>del (19 nt)</i>	MG272222
84	2	5	1	two 1 nt-insertion	209^210>ins^a (VH), 762^763>ins^t (VL)	MG272214
9	1	2	1	one 1 nt-substitution in VH, one 2nt-insertion + two 1 nt-substitution in VL	c322>t (VH), 658^659>ins^cc (VL), c659>t (VL), t660>a (VL)	MG272215
Total : 25				Total : 25		

722 ¹ Mutations are described according to the IMGT Scientific chart rules
723 (<http://www.imgt.org/IMGTScientificChart/Nomenclature/IMGTmutation.html>) (Lefranc, 1998)
724 (39). The mutations in the primers are shown in italics.

725

726

727 **Legends to Figures**

728 **Figure 1.** Human scFv-phagemid combinatorial library construction (A) and *in vivo* phage display
729 selection (B).

730

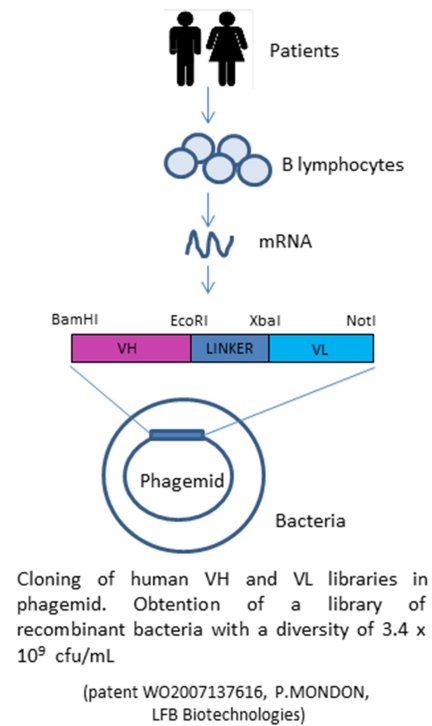
731 **Figure 2.** Obtaining recombinant scFv-phagemid bacteria from the AAR3 fraction, storage for NGS,
732 quantification and picking of individual clones.

733

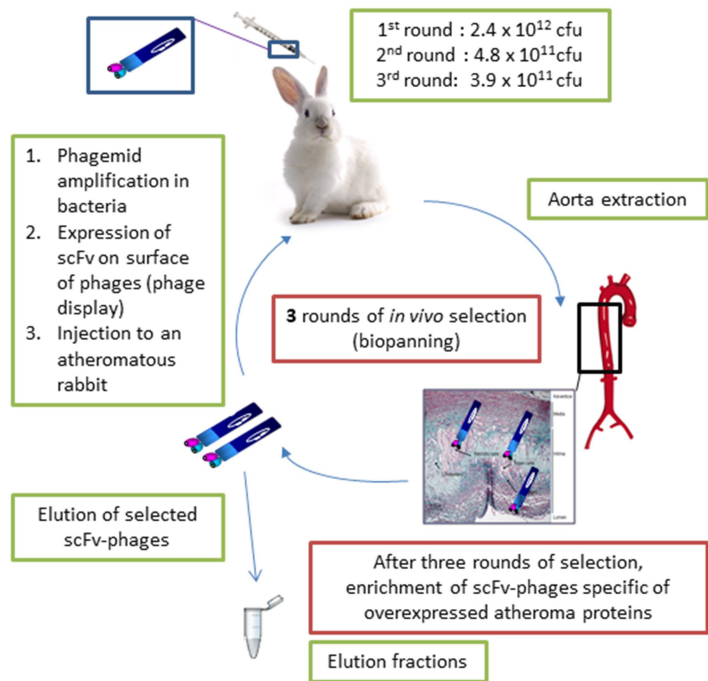
734 **Figure 3.** Primer design (A) and quality control pre- (B) and post-sequencing (C). (A) Primers
735 designed on the phagemid vector and used for scFv PCR amplification. The scFv (VH-LINKER-VL)
736 length range is between ~720 and ~800 bp (VH between ~350 and ~400 bp and VL between ~320
737 and ~350 bp). The linker is 53 bp including the EcoRI and XbaI sites. The PCR products are
738 expected to be ~1000 bp on average, including the 5' and 3' region and the primers. (B) Agarose gel
739 electrophoresis of PCR products. The DNA was amplified from the AAR3 fraction and PCR
740 products were analyzed on 1.2% (w/v) agarose gel. The band at ~1000bp corresponds to the expected
741 size for scFv amplicons. S1, S2, S3 and S4 correspond to the sample 1, 2, 3 and 4, respectively. The
742 Bioanalyzer trace of the four samples shows the purity of amplicons with a high-quality single peak.
743 (C) PacBio RS II CCS2 read length distribution using P6-C4 chemistry for one SMRT Cell (similar
744 results were obtained for the 15 SMRT Cells). Data are based on a one kb size-selected scFv library
745 using a 6h movie.

746

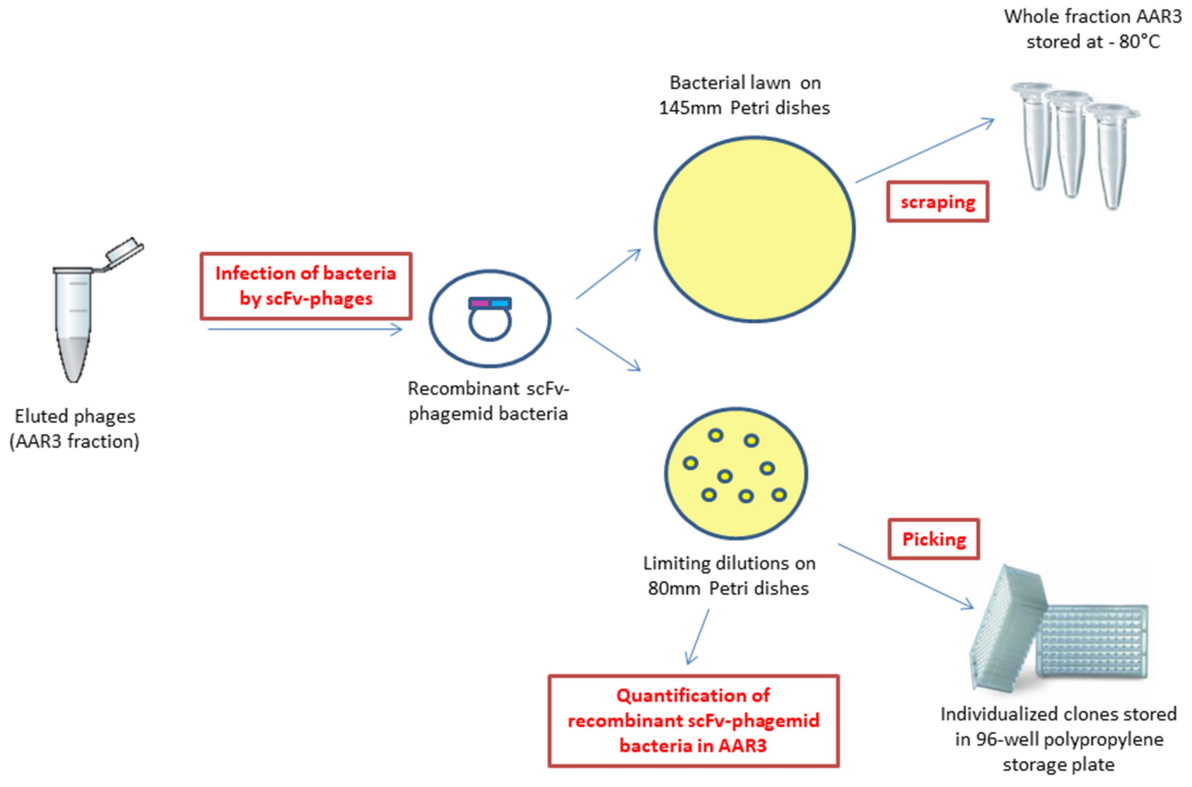
(A) Design of a fully human scFv library



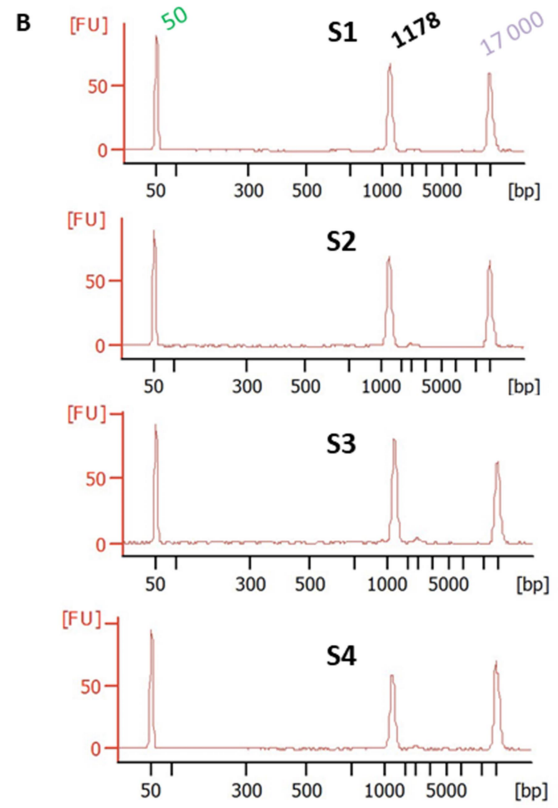
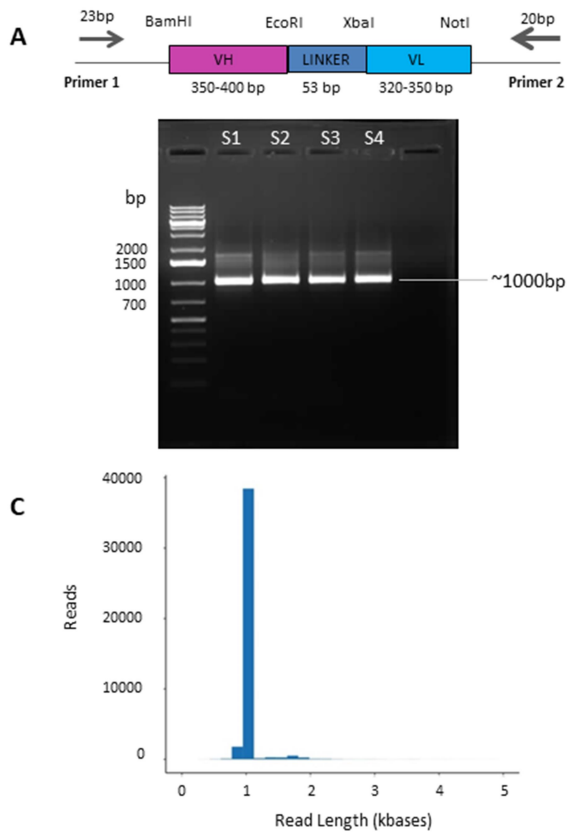
(B) Expression of ScFv on surface of phages (phage display) and three rounds of *in vivo* selection



747
748



749
750



751
 752

Supplementary Material

NGS PacBio full-length sequencing and IMGT/HighV-QUEST analysis of scFv from a phage-display combinatorial library

Audrey Hemadou, Véronique Giudicelli, Melissa Smith, Marie-Paule Lefranc*, Patrice Duroux, Sofia Kossida, Cheryl Heiner, John Kuijpers, Alexis Groppi, Jonas Korlach, Philippe Mondon, Florence Ottones, Marie-Josée Jacobin-Valat, Jeanny Laroche-Traineau and Gisèle Clofent-Sanchez*

*** Correspondence:**

Marie-Paule Lefranc Marie-Paule.Lefranc@igh.cnrs.fr

Gisèle Clofent-Sanchez gisele.clofent-sanchez@rmsb.u-bordeaux2.fr

Supplementary Tables

Table S1. Repartition of the AAR3 scFv, in the 15 SMRT Cells, based on their format (VH-VL or VL-VH, VH-VH, VL-VL) determined by IMGT/HighV-QUEST using the novel functionality for scFv.

PCR sample N°	Number of reads with 99.9% predicted accuracy	SMRT Cell N°	Number of PacBio CCS2 analysed reads	Number of filtered-in reads	Number of VH-VL or VL-VH	Number of VH-VH	Number of VL-VL
s1	91828	1	29,224	22,906	22819	13	74
		2	32,240	25,228	25136	14	78
		3	30,364	23,799	23723	8	68
s2	129640	4	34,082	26,657	26518	22	117
		5	33,510	26,407	26284	11	112
		6	31,980	25,032	24903	5	124
		7	30,068	23,695	23605	13	77
s3	115446	8	34,890	26,990	26817	15	158
		9	29,373	22,468	22343	7	118
		10	26,465	20,044	19959	6	79
		11	24,718	18,741	18597	12	132
s4	113644	12	25,128	19,120	19027	9	84
		13	23,693	17,756	17670	8	78
		14	32,293	24,762	24634	10	118
		15	32,530	25,054	24899	18	137
Total			450,558	348,659	346934	171	1554

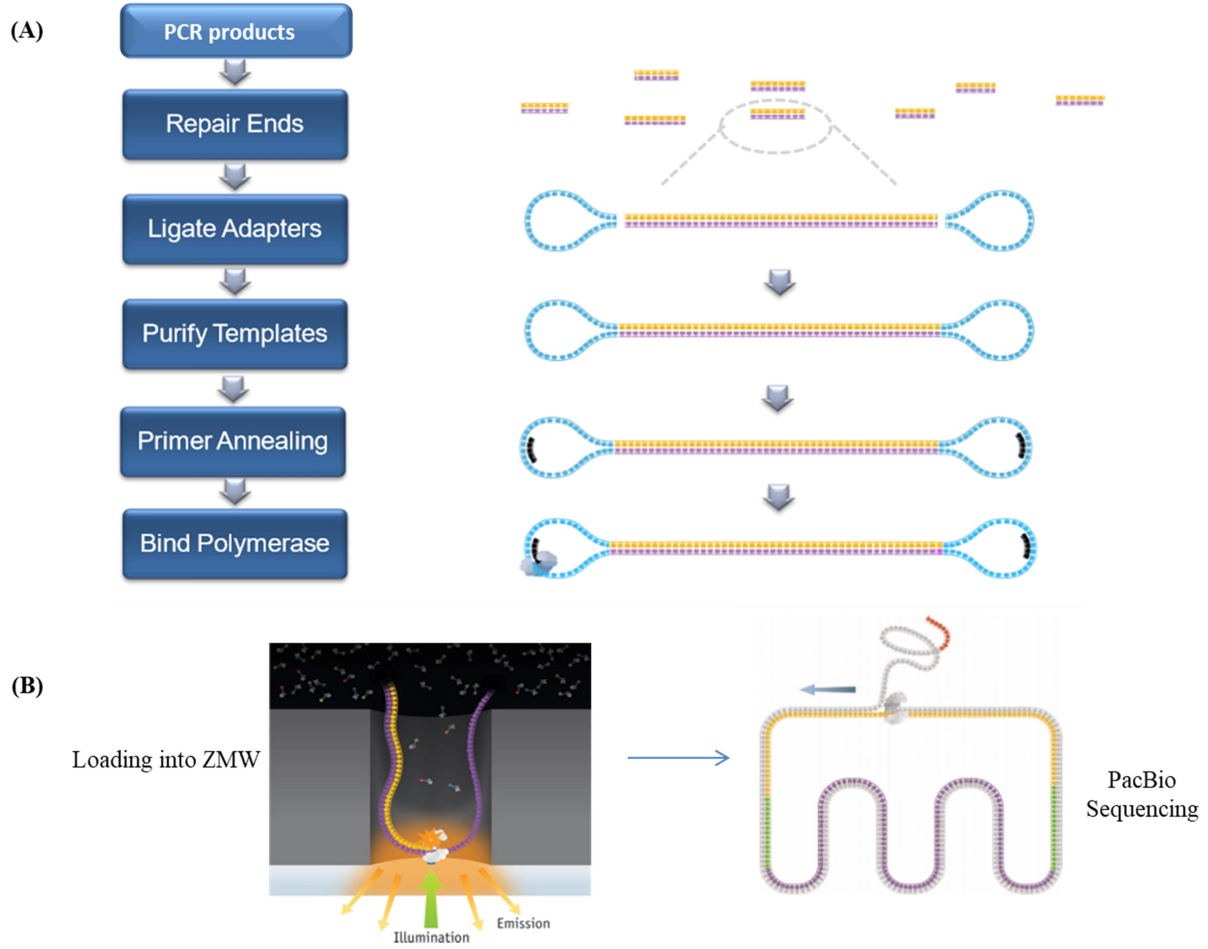
Figure S1. PacBio general workflow for amplicon sample preparation⁴ (A) and SMRTbell libraries loading and sequencing (B).

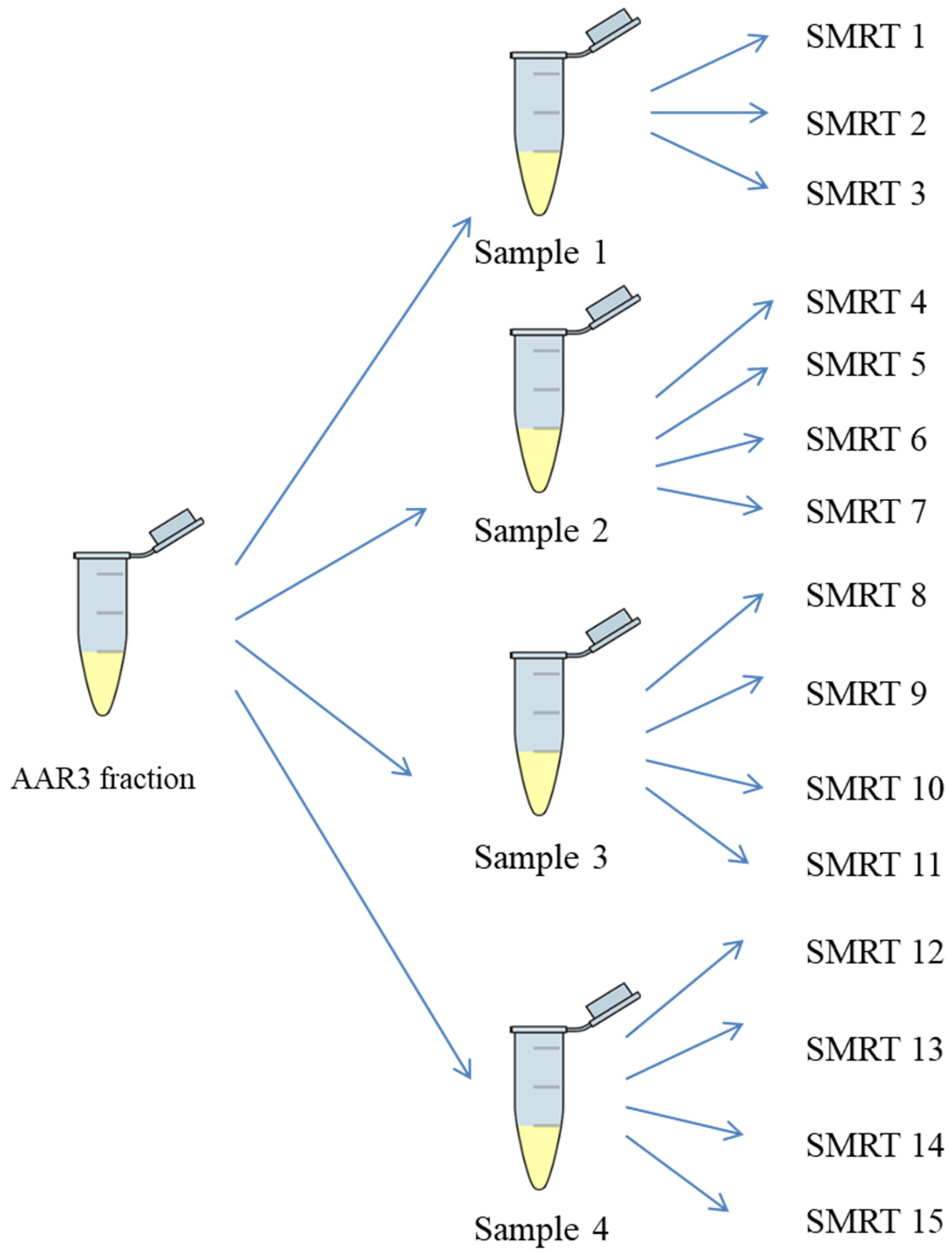
(<http://www.pacb.com/wp-content/uploads/2014/04/TemplatePreparation.pdf>)

Figure S2. Scheme of the four PCR amplifications of the AAR3 fraction and design of the corresponding 15 SMRT Cells. Sample 1 was loaded on 3 SMRT cells, and not 4, owing to limited sample quantity after library preparation.

Figure S3. Characterisation of P3 by IMGT/V-QUEST. (A) Parameters and positions of the two V domains VH and VL and of the linker. (B) Identification of the IMGT clonotype (AA) for VH domain (IGHV6-1*01 - IGHD6-13*01 - IGHJ4*02 [10.9.10]; CARQGSTYFDYW). The domain is unproductive because of a stop codon in FR2-IMGT position 52 according to the IMGT unique numbering (Lefranc et al. Dev. Comp. Immunol., 27, 55-77 (2003)). For biological analysis the stop codon was replaced with the tgg codon of the IGHV6-1*01. (C) Identification of the IMGT clonotype (AA) for VK domain: (IGKV1-39*01 - IGKJ4*01 [6.3.9]; CQQTYSAPPTF).

Figure S4. Alignment of the 25 P3-related reads with one of the P3 reads (among the fully identical on 977 bp). Dashes indicate alignment gaps. Numbers on the right indicate the number of nucleotides in the P3 sequence (excluding gaps created by the alignment). Primer FWD, VH, Linker, VL, Primer REV are delimited above the sequence. Mutations by comparison to P3 are highlighted in red, excepted those in the primers (only deletions) which are shown in pink. Large deletions at the 3' end of P3-related 58 and 81 are left in blank.







**A THANK YOU
for using IMGT/V-QUEST**

IMGT®, the international ImMunoGeneTics information system®

<http://www.imgt.org>

IMGT/V-QUEST program version: 3.4.8; IMGT/V-QUEST reference directory release: 201736-4

Species: **Homo sapiens**
 Receptor type or locus: **IG**
 IMGT directory reference set: **F+ORF+ in-frame P**
 Search for insertions and deletions: **yes**
 Analysis of scFv: **yes**

A. Detailed results for the IMGT/V-QUEST analysed sequences

Number of analysed sequences: 1
Number of analysed V-DOMAIN: 2

1 [P3_H](#), 2 [P3_K](#)

Identified scFv:

Sequence ID	5'V-DOMAIN ID	5'V-DOMAIN positions	5'V-DOMAIN length	linker positions	linker length	3'V-DOMAIN ID	3'V-DOMAIN positions	3'V-DOMAIN length
P3	1 P3_H	19..379	361	380..432	53	2 P3_K	433..754	322

This release of IMGT/V-QUEST uses [IMGT/JunctionAnalysis](#) for the analysis of the JUNCTION

Hyphens (-) show nucleotide identity, dots (.) represent gaps

B V-DOMAIN: 1 [P3_H](#) (associated V-DOMAIN: 2 [P3_K](#))

Sequence compared with the *Homo sapiens* (human) IG set from the [IMGT reference directory](#)

```
>P3_H
ggatccctgcaggtcgacacaggtacagctgcagcagtcaggtccaggactggtgaagccc
tcgcagacccctctcaactcaactgtgcaatctccggggacagtgctctagcaacaatttt
ggttggaaactggatcaggcagtcceccatcgagaggccttgagttagctgggaaggacatc
tacaggtccaagtggtataatgattatgcagtatctgtgagaagtccaataaccatcaac
tcagacacatccaagaaccagttctccctgcagctgaactctgactccggaggacacg
gcccgtgatctactgtcaagacagggcagcacttactctgactatggggccaggacc
atggccaactgtctcccaaatctcggggcgggtggctcggggcgggtgggtgggtctggggc
ggcggtctagagacatcgtgatgaccagctccatccctccctgtctgcatctgttagga
gacagagtcaccatcaactgcccggcaagtgcagacatagcagctctttagcttggttt
cagcagagaccgggagagaagccctaacctcctgatctatagtgcatccagtttgcaaat
ggggtccatcaaggttcagtgccagtggtctgggacagatttcaactcaccatcagc
agcttcaacctgaagatttgcaacttactcctgtcaacagacttacagtgcccctccc
acttccggcggaggaccagctggagatcaaatgcagggcggccgc
```

Result summary:	Unproductive IGH rearranged sequence (stop codons)		
V-GENE and allele	Homsap IGHV6-1*01 F	score = 1417	identity = 97.64% (290/297 nt)
J-GENE and allele	Homsap IGHJ4*02 F (a)	score = 186	identity = 87.50% (42/48 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD6-13*01 F	D-REGION is in reading frame 1	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.17.38.11]	[10.9.10]	CARQGSTYFDYW

Other possibilities: [Homsap_IGHJ2*01](#) (highest number of consecutive identical nucleotides)

C V-DOMAIN: 2 [P3_K](#) (associated V-DOMAIN: 1 [P3_H](#))

Sequence compared with the *Homo sapiens* (human) IG set from the [IMGT reference directory](#)

```
>P3_K
ggatccctgcaggtcgaccaggtacagctgcagcagtcaggtccaggactggtgaagccc
tcgcagacccctctcaactcaactgtgcaatctccggggacagtgctctagcaacaatttt
ggttggaaactggatcaggcagtcceccatcgagaggccttgagttagctgggaaggacatc
tacaggtccaagtggtataatgattatgcagtatctgtgagaagtccaataaccatcaac
ccagacacatccaagaaccagttctccctgcagctgaactctgactccggaggacacg
cccgtgtattactgtcaagacagggcagcacttactctgactatggggccaggacc
ctggtcaactgtctcctcagaattcgggtggcggtggctcggggcgggtgggtgggtctggggc
ggcggtctcagaacacatcgtgatgaccagctccatccctccctgtctgcatctgtagg
aacagagtcaccatcaacttccggggcaagtgcagaccatagcagctcttttagcttggttt
cagcagagaccgggagagaagccctaacctcctgatctatagtgcatccagtttgcaaat
ggggtccatcaaggttcagtgccagtggtctgggacagatttcaactcaccatcagc
agtctcaacctgaagatttgcaacttactcctgtcaacagacttacagtgcccctccc
acttccggcggaggaccagctggagatcaaatgcagggcggccgc
```

Result summary:	Productive IGK rearranged sequence: (no stop codon and in-frame junction)		
V-GENE and allele	Homsap IGKV1-39*01 F , or Homsap IGKV1D-39*01 F	score = 1237	identity = 93.91% (262/279 nt)
J-GENE and allele	Homsap IGKJ4*01 F 269	score = 157	identity = 94.29% (33/35 nt)
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[26.17.36.10]	[6.3.9]	CQQTYSAPPTF

```
-----Primer FWD-----|
1  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct 60
84  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
9   tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
85  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
40  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
7   tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
18  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
43  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
58  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
67  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
69  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
79  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
61  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
74  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
81  t caaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
5   tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
14  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
31  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
34  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
42  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
82  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
22  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
55  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
62  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
77  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
78  tgcaaattctatttcaaggagac agtcataatgaaatacctattgcctacggcagccgct
* *****
```

```
|<-----
1  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta 120
84  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
9   ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
85  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
40  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
7   ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
18  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
43  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
58  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
67  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
69  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
79  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
61  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
74  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
81  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
5   ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
14  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
31  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
34  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
42  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
82  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
22  ggattggtattactcgcggcccagccggccatggcgggatccctgcaggtcgac caggta
```

55 ggattggttattactcgcgggccagccggccatggcgggatccctgcaggtcgac caggta
62 ggattggttattactcgcgggccagccggccatggcgggatccctgcaggtcgac caggta
77 ggattggttattactcgcgggccagccggccatggcgggatccctgcaggtcgac caggta
78 ggattggttattactcgcgggccagccggccatggcgggatccctgcaggtcgac caggta

1 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt 180
84 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
9 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
85 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
40 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
7 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
18 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
43 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
58 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
67 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
69 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
79 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
61 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
74 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
81 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
5 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
14 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
31 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
34 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
42 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
82 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
22 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
55 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
62 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
77 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt
78 cagctgcagcagtcaggtccaggactggggaagccctcgagaccctctcactcacctgt

1 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc 239
84 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
9 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
85 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
40 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
7 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
18 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
43 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
58 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
67 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
69 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
79 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
61 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
74 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
81 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
5 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
14 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc
31 gccatctccggggacagtggtctctagcaa-caattttggttggaaactggatcaggcagtc

34 gccatctccggggacagtgctctctagcaa-caat tttgggtggaactggatcaggcagtc
42 gccatctccggggacagtgctctctagcaa-caat tttgggtggaactggatcaggcagtc
82 gccatctccggggacagtgctctctagcaa-caat tttgggtggaactggatcaggcagtc
22 gccatctccggggacagtgctctctagcaa-caat tttgggtggaactggatcaggcagtc
55 gccatctccggggacagtgctctctagcaa-caat tttgggtggaactggatcaggcagtc
62 gccatctccggggacagtgctctctagcaa-caat tttgggtggaactggatcaggcagtc
77 gccatctccggggacagtgctctctagcaa-caat tttgggtggaactggatcaggcagtc
78 gccatctccggggacagtgctctctagcaa-caat tttgggtggaactggatcaggcagtc

1 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga 299
84 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
9 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
85 cc[atcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
40 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
7 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
18 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
43 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
58 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
67 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
69 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
79 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
61 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
74 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
81 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
5 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
14 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
31 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
34 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
42 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
82 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
22 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
55 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
62 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
77 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
78 cccatcgagagggccttgagtagctgggaaggacatactacaggtccaagtgggtataatga
** *****

----- VH -----
1 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt 359
84 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt
9 ttatgcagtatctgtgagaagtc[gaataaaccatcaaccagacacatccaagaaccagtt
85 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt
40 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt
7 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt
18 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt
43 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt
58 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt
67 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt
69 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt
79 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt
61 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt
74 ttatgcagtatctgtgagaagtcgaataaaccatcaaccagacacatccaagaaccagtt

81 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt
5 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt
14 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt
31 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt
34 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt
42 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt
82 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt
22 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt
55 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt
62 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt
77 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt
78 ttatgcagtatctgtgagaagtcgaataaacatcaaccagacacatccaagaaccagtt

1 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca 419
84 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
9 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
85 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
40 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
7 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
18 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
43 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
58 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
67 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
69 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
79 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
61 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
74 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
81 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
5 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
14 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
31 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
34 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
42 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
82 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
22 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
55 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
62 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
77 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca
78 ctccctgcagctgaactctgtgactcccaggacacggccgtgtattactgtgcaagaca

----->|<-----
1 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt 479
84 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
9 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
85 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
40 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
7 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
18 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
43 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
58 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
67 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt

69 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 79 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 61 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 74 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 81 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 5 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 14 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 31 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 34 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 42 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 82 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 22 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 55 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 62 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 77 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt
 78 gggcagcacttacttcgactattggggccagggcaccctggtcactgtctcctcag aatt

----- Linker -----> | <-----

1 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat 539
 84 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 9 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 85 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 40 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 7 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 18 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 43 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 58 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 67 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 69 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 79 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 61 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 74 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 81 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 5 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 14 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 31 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 34 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 42 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 82 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 22 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 55 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 62 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 77 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat
 78 cggtggcggtggctcgggcggtgggtgggtctgggtggcggttctaga gacatcgtgat

1 gaccagtcctccatcctcctgtctgcatctgtaggagacagagtcaccatcacttgccg 599
 84 gaccagtcctccatcctcctgtctgcatctgtaggagacagagtcaccatcacttgccg
 9 gaccagtcctccatcctcctgtctgcatctgtaggagacagagtcaccatcacttgccg
 85 gaccagtcctccatcctcctgtctgcatctgtaggagacagagtcaccatcacttgccg
 40 gaccagtcctccatcctcctgtctgcatctgtaggagacagagtcaccatcacttgccg

7 gaccgggtctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
18 gaccgggtctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
43 gaccgggtctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
58 gaccgggtctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
67 gaccgggtctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
69 gaccgggtctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
79 gaccgggtctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
61 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
74 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
81 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
5 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
14 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
31 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
34 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
42 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
82 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
22 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
55 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
62 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
77 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg
78 gaccagtgctccatcctccctgtctgcatctgtaggagacagagtcaccatcacttgccg

1 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc- 658
84 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
9 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
85 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
40 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
7 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
18 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
43 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
58 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
67 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
69 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
79 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
61 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
74 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
81 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
5 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
14 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
31 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
34 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
42 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
82 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
22 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
55 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
62 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
77 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
78 ggcaagtcagaccattagcagctcttttagcttggtttcagcagagaccgggagaagccc-
***** ** *****

VL
1 -ctaacctcctgatctatagtgcattcagtttgcaaagtgggtcccattcaaggttcagt 717

84 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
9 ctaaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
85 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
40 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
7 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
18 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
43 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
58 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
67 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
69 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
79 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
61 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
74 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
81 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
5 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
14 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
31 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
34 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
42 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
82 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
22 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
55 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
62 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
77 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt
78 -ctaacctcctgatctatagtgcacccagtttgcaaagtggggtcccatcaaggttcagt

1 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt 776
84 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
9 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
85 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
40 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
7 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
18 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
43 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
58 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
67 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
69 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
79 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
61 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
74 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
81 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
5 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
14 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
31 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
34 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
42 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
82 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
22 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
55 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
62 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
77 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt
78 ggcagtggtctctgggacagatttcactctcaccatcagcagtcctt-caacctgaagatt

** ***** **

1 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa 836
84 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
9 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
85 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
40 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
7 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
18 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
43 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
58 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
67 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
69 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
79 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
61 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
74 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
81 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
5 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
14 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
31 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
34 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
42 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
82 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
22 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
55 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
62 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
77 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa
78 tgcaacttactcctgtcaacagacttacagtgccctcccactttcggcggagggaccaa

----->|
1 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg 896
84 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
9 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
85 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
40 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
7 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
18 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
43 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
58 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacggg-----
67 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
69 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
79 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
61 gctggatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
74 gctggatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
81 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
5 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
14 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
31 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
34 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
42 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
82 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg
22 gctggagatcaaag tcgaggcggccgcgatagagaacaacggcaacgggaacggcacagg

55 gctggagatcaaag tcgagggcggccgcatagagaaacaacggcaacgggaacggcacagg
 62 gctggagatcaaag tcgagggcggccgcatagagaaacaacggcaacgggaacggcacagg
 77 gctggagatcaaag tcgagggcggccgcatagagaaacaacggcaacgggaacggcacagg
 78 gctggagatcaaag tcgagggcggccgcatagagaaacaacggcaacgggaacggcacagg
 **** * 9

1 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc 956
 84 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 9 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 85 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 40 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 7 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 18 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 43 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 58 -----
 67 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 69 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 79 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 61 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 74 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 81 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 5 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 14 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 31 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 34 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 42 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 82 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 22 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 55 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 62 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 77 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc
 78 agatgttttgcttaacgttctttcgaagtgtgcccggctctaaaaaattcaggcctgctcc

| ←--- Primer REV ---→
 1 c gccgctgcttttgcacgtga 977
 84 c gccgctgcttttgcacgtga
 9 c gccgctgcttttgcacgtga
 85 c gccgctgcttttgcacgtga
 40 c gccgctgcttttgcacgtga
 7 c gccgctgcttttgcacgtga
 18 c gccgctgcttttgcacgtga
 43 c gccgctgcttttgcacgtga
 58 -----
 67 c gccgctgcttttgcacgtga
 69 c gccgctgcttttgcacgtga
 79 c gccgctgcttttgcacgtga
 61 c gccgctgcttttgcacgtga
 74 c gccgctgcttttgcacgtg
 81 c g-----
 5 c gccgctgcttttgcacgtga
 14 c gccgctgcttttgcacgtga
 31 c gccgctgcttttgcacgtga
 34 c gccgctgcttttgcacgtga

42 c gccgctgcttttgcacgtga
82 c gccgctgcttttgcacgtga
22 c gccgctgcttttgcacgtga
55 c gccgctgcttttgcacgtga
62 c gccgctgcttttgcacgtga
77 c gccgctgcttttgcacgtga
78 c gccgctgcttttgcacg

Bibliographie

1. White, S. J., Nicklin, S. A., Sawamura, T. & Baker, A. H. Identification of peptides that target the endothelial cell-specific LOX-1 receptor. *Hypertens. Dallas Tex* 1979 **37**, 449–455 (2001).
2. Molenaar, T. J. M. *et al.* P-selectin as a candidate target in atherosclerosis. *Biochem. Pharmacol.* **66**, 859–866 (2003).
3. Luque, M. C. A. *et al.* Phage display identification of CD100 in human atherosclerotic plaque macrophages and foam cells. *PloS One* **8**, e75772 (2013).
4. Cooksley-Decasper, S. *et al.* Antibody Phage Display Assisted Identification of Junction Plakoglobin as a Potential Biomarker for Atherosclerosis. *PLOS ONE* **7**, e47985 (2012).
5. Huang, Y. *et al.* An abundant dysfunctional apolipoprotein A1 in human atheroma. *Nat. Med.* **20**, 193–203 (2014).
6. Kumano-Kuramochi, M., Fujimura, T., Komba, S., Maeda-Yamamoto, M. & Machida, S. Screening, expression, and characterization of an anti-human oxidized low-density lipoprotein single-chain variable fragment. *J. Biosci. Bioeng.* **122**, 287–293 (2016).
7. Robert, R. *et al.* Identification of human scFvs targeting atherosclerotic lesions: selection by single round in vivo phage display. *J. Biol. Chem.* **281**, 40135–40143 (2006).
8. Deramchia, K. *et al.* In Vivo Phage Display to Identify New Human Antibody Fragments Homing to Atherosclerotic Endothelial and Subendothelial Tissues. *Am. J. Pathol.* **180**, 2576–2589 (2012).
9. Mondon, P. *et al.* Method for generation of human hyperdiversified antibody fragment library. *Biotechnol. J.* **2**, 76–82 (2007).
10. Rekhter, M. D. *et al.* Animal model that mimics atherosclerotic plaque rupture. *Circ. Res.* **83**, 705–713 (1998).
11. Moldavan, A. PHOTO-ELECTRIC TECHNIQUE FOR THE COUNTING OF MICROSCOPICAL CELLS. *Science* **80**, 188–189 (1934).

12. Barter, P. J. *et al.* Antiinflammatory Properties of HDL. *Circ. Res.* **95**, 764–772 (2004).
13. Emini Veseli, B. *et al.* Animal models of atherosclerosis. *Eur. J. Pharmacol.* (2017). doi:10.1016/j.ejphar.2017.05.010
14. Zhao, C. F. & Herrington, D. M. The function of cathepsins B, D, and X in atherosclerosis. *Am. J. Cardiovasc. Dis.* **6**, 163–170 (2016).
15. Zhao, B. *et al.* Macrophage-specific transgenic expression of cholesteryl ester hydrolase significantly reduces atherosclerosis and lesion necrosis in Ldlr^{-/-} mice. *J. Clin. Invest.* **117**, 2983–2992 (2007).
16. Miyazaki, A. *et al.* Expression of ACAT-1 protein in human atherosclerotic lesions and cultured human monocytes-macrophages. *Arterioscler. Thromb. Vasc. Biol.* **18**, 1568–1574 (1998).
17. Pang, Q. *et al.* UFM1 Protects Macrophages from oxLDL-Induced Foam Cell Formation Through a Liver X Receptor α Dependent Pathway. *J. Atheroscler. Thromb.* **22**, 1124–1140 (2015).
18. Yazgan, B., Ustunsoy, S., Karademir, B. & Kartal-Ozer, N. CD36 as a biomarker of atherosclerosis. *Free Radic. Biol. Med.* **75 Suppl 1**, S10 (2014).
19. Furuhashi, M., Saitoh, S., Shimamoto, K. & Miura, T. Fatty Acid-Binding Protein 4 (FABP4): Pathophysiological Insights and Potent Clinical Biomarker of Metabolic and Cardiovascular Diseases. *Clin. Med. Insights Cardiol.* **8**, 23–33 (2014).
20. Yancey, P. G. *et al.* Low-density lipoprotein receptor-related protein 1 prevents early atherosclerosis by limiting lesional apoptosis and inflammatory Ly-6Chigh monocytosis: evidence that the effects are not apolipoprotein E dependent. *Circulation* **124**, 454–464 (2011).
21. Fernández-Hernando, C. *et al.* Genetic evidence supporting a critical role of endothelial caveolin-1 during the progression of atherosclerosis. *Cell Metab.* **10**, 48–54 (2009).

22. de Gaetano, M., Dempsey, E., Marcone, S., James, W. G. & Belton, O. Conjugated linoleic acid targets β 2 integrin expression to suppress monocyte adhesion. *J. Immunol. Baltim. Md 1950* **191**, 4326–4336 (2013).
23. Pothula, A., Serebruany, V. L., Gurbel, P. A., McKenzie, M. E. & Atar, D. Pathophysiology and therapeutic modification of thrombin generation in patients with coronary artery disease. *Eur. J. Pharmacol.* **402**, 1–10 (2000).
24. Al-Ansari, S., Zeebregts, C. J., Slart, R. H. J. A., Peppelenbosch, M. & Tio, R. A. Galectins in atherosclerotic disease. *Trends Cardiovasc. Med.* **19**, 164–169 (2009).
25. Hogas, S. *et al.* Potential novel biomarkers of cardiovascular dysfunction and disease: cardiostrophin-1, adipokines and galectin-3. *Arch. Med. Sci. AMS* **13**, 897–913 (2017).
26. He, B., Huang, B., Lu, Z., He, W. & Jiang, H. Galectin-3: A potential new target for upstream therapy of atrial fibrillation. *Int. J. Cardiol.* **203**, 1131–1132 (2016).
27. Falcone, C. *et al.* Galectin-3 plasma levels and coronary artery disease: a new possible biomarker of acute coronary syndrome. *Int. J. Immunopathol. Pharmacol.* **24**, 905–913 (2011).
28. Rebholz, C. M. *et al.* Plasma galectin-3 levels are associated with the risk of incident chronic kidney disease. *Kidney Int.* (2017). doi:10.1016/j.kint.2017.06.028
29. Tao, L. *et al.* Galectin-3 Expression in Colorectal Cancer and its Correlation with Clinical Pathological Characteristics and Prognosis. *Open Med. Wars. Pol.* **12**, 226–230 (2017).
30. D'Alessandria, C. *et al.* Noninvasive In Vivo Imaging and Biologic Characterization of Thyroid Tumors by ImmunoPET Targeting of Galectin-3. *Cancer Res.* **76**, 3583–3592 (2016).
31. Xue, J. *et al.* Galectin-3 promotes caspase-independent cell death of HIV-1-infected macrophages. *FEBS J.* **284**, 97–113 (2017).

32. Martinez, R., Burrage, A. M., Wiethoff, C. M. & Wodrich, H. High temporal resolution imaging reveals endosomal membrane penetration and escape of adenoviruses in real time. *Methods Mol. Biol. Clifton NJ* **1064**, 211–226 (2013).
33. Takemoto, Y. *et al.* Galectin-3 Regulates Atrial Fibrillation Remodeling and Predicts Catheter Ablation Outcomes. *JACC Basic Transl. Sci.* **1**, 143–154 (2016).
34. Gawor, M. *et al.* The usefulness of soluble ST2 and galectin-3 as novel biomarkers for better risk stratification in hypertrophic cardiomyopathy. *Kardiol. Pol.* (2017). doi:10.5603/KP.a2017.0118
35. Zhu, W. *et al.* The role of galectin-3 in endocytosis of advanced glycation end products and modified low density lipoproteins. *Biochem. Biophys. Res. Commun.* **280**, 1183–1188 (2001).
36. Pugliese, G., Iacobini, C., Blasetti Fantauzzi, C. & Menini, S. The dark and bright side of atherosclerotic calcification. *Atherosclerosis* **238**, 220–230 (2015).
37. Ortega, N., Behonick, D. J., Colnot, C., Cooper, D. N. W. & Werb, Z. Galectin-3 Is a Downstream Regulator of Matrix Metalloproteinase-9 Function during Endochondral Bone Formation. *Mol. Biol. Cell* **16**, 3028–3039 (2005).
38. Silverman, A. M., Nakata, R., Shimada, H., Sposto, R. & DeClerck, Y. A. A GALECTIN-3-DEPENDENT PATHWAY UPREGULATES INTERLEUKIN-6 IN THE MICROENVIRONMENT OF HUMAN NEUROBLASTOMA. *Cancer Res.* **72**, (2012).
39. Friedrichs, J., Manninen, A., Muller, D. J. & Helenius, J. Galectin-3 regulates integrin alpha2beta1-mediated adhesion to collagen-I and -IV. *J. Biol. Chem.* **283**, 32264–32272 (2008).
40. Van, F. den B., Buicu, C., Sobel, M. E., Liu, F. T. & Castronovo, V. Galectin-3, a laminin binding protein, fails to modulate adhesion of human melanoma cells to laminin. *Neoplasia* **42**, 215–219 (1995).
41. Lagana, A. *et al.* Galectin binding to Mgat5-modified N-glycans regulates fibronectin matrix remodeling in tumor cells. *Mol. Cell. Biol.* **26**, 3181–3193 (2006).

42. Reddy, S. T. *et al.* Monoclonal antibodies isolated without screening by analyzing the variable-gene repertoire of plasma cells. *Nat. Biotechnol.* **28**, 965–969 (2010).
43. Naso, M. F., Lu, J. & Panavas, T. Deep sequencing approaches to antibody discovery. *Curr. Drug Discov. Technol.* **11**, 85–95 (2014).
44. Matochko, W. L. *et al.* Deep sequencing analysis of phage libraries using Illumina platform. *Methods San Diego Calif* **58**, 47–55 (2012).
45. Ravn, U. *et al.* By-passing in vitro screening--next generation sequencing technologies applied to antibody display and in silico candidate selection. *Nucleic Acids Res.* **38**, e193 (2010).

Conclusion générale et perspectives

Sélection d'anticorps humains et découverte de nouveaux biomarqueurs

Les plaques d'athérome ont été décrites il y a de cela presque trois siècles par le docteur allemand Friedrich KRELL. Ce premier, ayant découvert des concrétions calciques au niveau des artères, a nommé ces plaques, des plaques osseuses. Ce n'est qu'en 1933 que le terme d'athérosclérose fût utilisé par le docteur Jean Frédéric Martin LOBSTEIN, médecin strasbourgeois, pour décrire le durcissement de la paroi artérielle. Depuis, la connaissance autour de cette pathologie est en constante évolution. Il est clair aujourd'hui que cette pathologie complexe fait intervenir une multitude de cellules immunes et non immunes ainsi que de nombreux facteurs moléculaires, l'initiation étant provoquée par une accumulation de lipides oxydés au sein de la paroi. Malgré l'évolution des connaissances, le diagnostic des patients à risque reste problématique. Il n'existe aujourd'hui aucun moyen de diagnostic non invasif permettant de distinguer les plaques instables des plaques stables. En effet, il a été montré que plus que le rétrécissement de la lumière du vaisseau, la composition de la plaque influence grandement sa dangerosité liée au risque de rupture et par voie de conséquence les problèmes cardiovasculaires qui s'ensuivent. Malgré les effets bénéfiques des traitements prescrits comme les statines, afin de réduire le taux de cholestérol dans le sang, très peu de biomarqueurs sont exploités pour le développement d'outils de diagnostic et de thérapie de la plaque d'athérome. C'est donc dans l'optique de faire évoluer ces techniques que la recherche se tourne vers la découverte de nouvelles cibles pour le ciblage de l'athérome.

Comme démontré dans les résultats, l'approche par phage display *in vivo* serait parfaitement adaptée. Si on considère les problèmes de reconnaissance de cibles *in vivo* rencontrés après les sélections des ligands *in vitro*, les sélections *in vivo* s'imposent alors comme des alternatives permettant d'accéder à de nombreuses protéines dans leur conformation native ainsi que sous l'influence de leur microenvironnement. L'intérêt est donc ici d'interroger la plaque grâce à des phage-scFv issus d'une banque combinatoire humaine hyperdiversifiée afin de sélectionner des couples anticorps/cible.

Suite à la sélection *in vivo*, nous avons, ici, réussi à sélectionner un pool de 200 anticorps par cytométrie de flux à moyen débit, technique implémentée au cours de cette thèse. L'un de ces anticorps a, par ailleurs, été caractérisé jusqu'à la découverte de la galectine 3 en tant que cible grâce à la spectrométrie de masse et des tests de réactivité *in vitro*. Parallèlement au criblage *in vivo* puis *in vitro* par cytométrie de flux, nous avons montré également qu'il était possible d'aller « repêcher » des anticorps enrichis lors de la sélection *in vivo* par les dernières techniques de séquençage haut débit. Par la suite, il est possible de synthétiser ces clones issus d'études *in silico* par synthèse chimique de gènes sur mesure appelé aussi « rescue », pour tester *in fine* leur réactivité *in vitro*.

L'une des perspectives immédiates est donc d'identifier les protéines cibles des anticorps pour lesquels on a démontré une forte bioréactivité *in vitro* par cytométrie de flux ainsi qu'une reconnaissance *ex vivo* des sections de plaques d'athérome (C10PL20, E10PL25, C4PL23 et B7PL27). Pourra être utilisée la technologie d'immunoprécipitation suivie d'une étude en protéomique par spectrométrie de masse mise en place au cours de cette thèse. Quelques difficultés ont cependant été rencontrées pour la détermination de la cible des autres anticorps à fort potentiel biologique. L'étude protéomique sur les protéines athéromateuses de lapin

nous ayant par ailleurs permis de mettre en exergue de nombreuses protéines sur-représentées dans les échantillons, il peut être avantageux, en termes de rapidité et d'efficacité, d'envisager des techniques d'identification plus systématiques comme la technologie des puces à protéines qui pourrait nous permettre de nous affranchir des contraintes liées à l'immunoprécipitation. La plateforme GeT-Biopuces, à l'INSA à Toulouse, propose de fabriquer des puces à protéines à façon. Ainsi, nous pourrions envisager une collaboration pour un « design » de biopuces avec des protéines spécifiques de l'inflammation pour pouvoir ensuite cribler nos anticorps couplés à des fluorochromes pour détecter la fixation.

Il peut être aussi envisagé d'étendre les méthodologies mises en place, à savoir par étude de bio-réactivité ou par étude *in silico*, à la caractérisation d'autres anticorps issus de la sélection *in vivo* (**Figure 61**). Si l'on souhaite poursuivre l'étude *in silico*, il faudra alors procéder à un « rescue » des séquences obtenues par NGS. Il est à noter que seulement deux des fractions issues de la sélection *in vivo* ont été séquencées à l'heure actuelle. Il serait intéressant de procéder au séquençage des différentes fractions pour (1) avoir une idée de l'enrichissement qui s'est produit durant la sélection *in vivo* et (2) d'identifier les associations VH-VL les plus représentées.

Ces deux études parallèles pourront ainsi nous permettre d'obtenir un panel d'anticorps dont les cibles pourront être déterminées. Ainsi, nous aurons accès à un éventail de biomarqueurs afin d'avoir une meilleure compréhension de l'athérogénèse.

Il est important de garder en mémoire que ces deux méthodes parallèles de criblage sont complémentaires. En effet, le séquençage haut débit peut nous permettre de sélectionner des anticorps enrichis durant la sélection, comme le clone P3, retrouvé dans les données NGS. Cependant, les criblages *in vitro* ont également mené à la sélection de clones, comme l'anticorps C10PL20 présentant une forte réactivité, qui n'ont pas été retrouvés dans les données de séquençage en tant qu'association VH-VL les plus représentées. Les deux criblages sont donc intéressants : l'un pour la sélection de clones enrichis, l'autre pour la sélection de clones qui peuvent être sous-représentés mais avoir un intérêt biologique non négligeable.

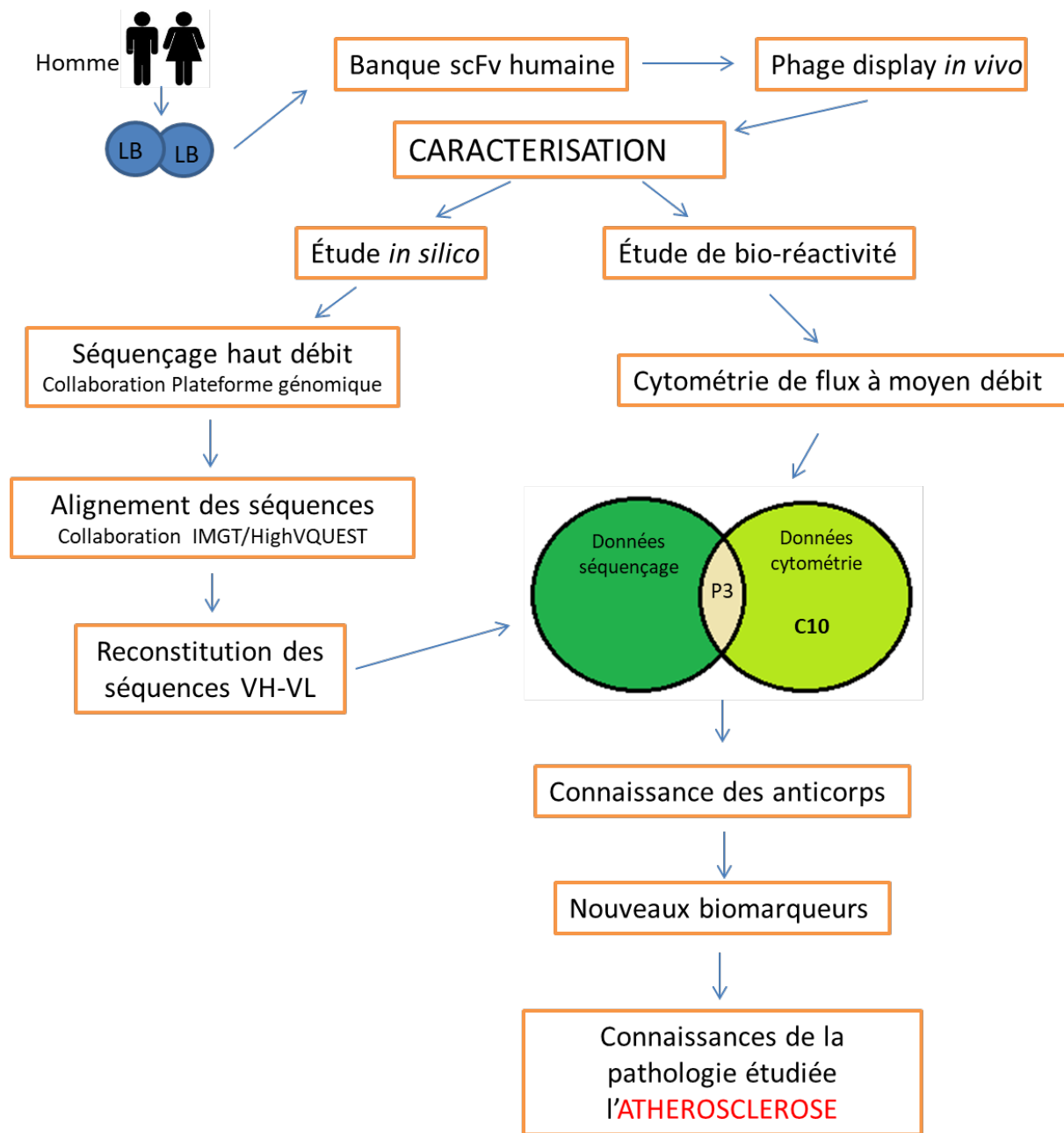


Figure 61 Processus de sélection d'anticorps humains par deux méthodes *in silico* et *in vitro* pour aboutir à la découverte de nouveaux biomarqueurs et à une meilleure compréhension de l'athérosclérose

Couplage des anticorps humains à des nanoparticules pour une imagerie non-invasive de la plaque

Le but ultime de ces recherches est le couplage des anticorps découverts, comme l'anticorps P3 mais également tous ceux présentés dans la partie « Résultats », à des agents de contrastes ou à des radioéléments. Ces derniers, ainsi fonctionnalisés, pourront être utilisés pour la visualisation de la plaque en imagerie non invasive. L'idée est de pouvoir diagnostiquer ces plaques avant qu'elles ne soient symptomatiques et que l'évènement cardio-vasculaire ne se produise. Plusieurs agents de contrastes en IRM et certains radioéléments pour l'imagerie TEP ou SPECT peuvent être utilisés. Les anticorps sélectionnés et en particulier les anticorps humains, sont un vecteur de choix de par leur spécificité d'une part mais également leur immunogénicité moindre. Ce dernier est un facteur très important, car comme discuté tout au long de ce manuscrit, ces anticorps humains pourront être transférés directement en clinique en réduisant les risques de complications mais également éviter leur élimination avant qu'ils n'aient pu se fixer sur leur cible. Grâce à l'ingénierie des anticorps, plusieurs formats pourront être également synthétisés (scFv, scFv-Fc, IgG1...)

Comme nous l'avons vu, il existe différents types d'agents de contrastes utilisés en IRM. Le gadolinium, qui est un agent paramagnétique, influence le T1 des molécules d'eau qui se situent à proximité, induisant un hypersignal. Ces agents, comme décrits dans la littérature, peuvent être toxiques et doivent donc être complexés avec du DTPA (diéthylène-triamine-penta-acétate) ou du DOTA (tétra-azacyclododécane-tétra-acétate) pour pouvoir être utilisés *in vivo*. A contrario, l'oxyde de fer, que nous utilisons au laboratoire en collaboration avec des chimistes pour le développement d'agents de contrastes, est un agent super-paramagnétique. Il diminue les temps de relaxation en T2 et génère un hyposignal. Un des avantages de l'oxyde de fer est qu'il peut être éliminé par la voie du métabolisme du fer, présente dans l'organisme. Plusieurs types de particules contenant de l'oxyde de Fer existent : les SPIO (superParamagnetic IronOxide), les USPIO (Ultra SuperParamagnetic IronOxide), les CLIO (CrossLinked IronOxide), les VUSPIO (Versatile UltraSmall SuperParamagnetic IronOxide) ou encore les SLN (Solid lipid nanoparticle).

Les protocoles de couplage des anticorps dépendront essentiellement de l'agent de contraste ou du radioélément considéré. Les VUSPIO, par exemple, portent à leur surface des macromolécules de PEG-bNH₂ (polyoxyde d'éthylène glycol bis-amine) qui permettent la fixation des fonctions thiol des anticorps. Des vecteurs plasmidiques ont été construits pour permettre l'expression d'anticorps possédant 2 cystéines en bout de chaîne constante dans le système baculovirus/cellules d'insectes (*Sf9*) (BaculoMab) ou dans le système levure (*Pischia Pastoris*). Ainsi, une réduction ménagée au TCEP (tris(2-carboxyethyl)phosphine) permettra de réduire les fonctions thiol (SH) conduisant à la liaison de ces fonctions SH libres au PEG-bNH₂^{1,2}. Il est à noter que des tests de fonctionnalisation ont été faits à l'ICMCB pour le couplage des scFv à des VUSPIO. Une optimisation du protocole a permis de fixer 14 fragments scFv à la surface des VUSPIO. Les résultats de Biacore (test d'affinité et d'avidité) ont permis de mettre en évidence que plus l'on augmente le nombre d'anticorps à la surface des nanoparticules plus l'avidité est importante vis-à-vis d'une protéine d'intérêt (Larivière et al. In progress). Un autre type de couplage a été envisagé, notamment par l'ajout, toujours dans le système BaculoMab, d'un motif LPETG en fin de chaîne constante qui, grâce à la sortase, peut être lié à une queue PEG-polyG³.

Dans un futur proche, ce système de nanoparticules fonctionnalisées pourra également servir pour la délivrance de drogues *in situ*. En tant que preuve de concept, nous avons montré que de la prostacycline (PGI₂) encapsidée dans des SLN, fabriqués par P. Barthélémy, INSERM U1212, gardait son rôle d'antiagrégant plaquettaire. En effet, l'agrégation du PRP (Platelet rich plasma) induite par ajout d'ADP (Adénosine diphosphate) ou de TRAP-6 (Peptide-6 activateur de récepteur de thrombine) est inhibée après incubation avec des SLN chargées en PGI₂ (Figure 62 + article joint en annexe)⁴.

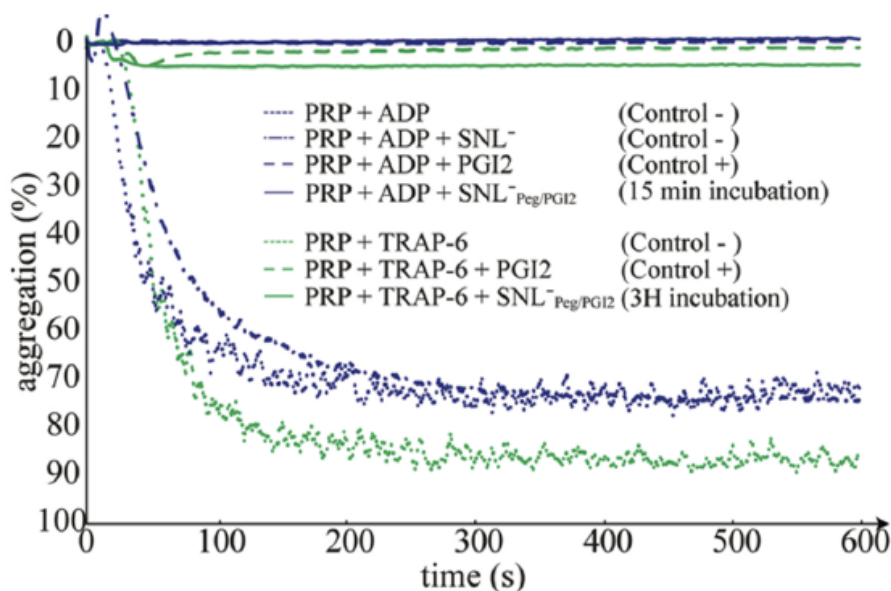


Figure 62 Inhibition de l'agrégation plaquettaire par des SLN chargées en PGI₂. Le pourcentage d'agrégation a été suivi en fonction du temps. Deux cent sept microlitres de PRP (207 μ L), sous agitation dans des cuvettes à 37°C, ont été mis en présence d'ADP (courbes bleues) ou de TRAP (courbes vertes) pour favoriser l'agrégation plaquettaire. Les SLN⁻_{Peg / PGI₂} (lignes pleines) ainsi que de la PGI₂ libre inhibent totalement l'agrégation plaquettaire induite à la fois par l'ADP (15 min d'incubation) ou par le TRAP-6 (incubation 3 h), alors que les nanoparticules SLN⁻ comme contrôle négatif (ligne en point-tiret) n'ont aucun effet sur l'agrégation induite par l'ADP.

Bibliographie

1. Mornet, S., Portier, J. & Duguet, E. A method for synthesis and functionalization of ultrasmall superparamagnetic covalent carriers based on maghemite and dextran. *J. Magn. Magn. Mater.* **293**, n° **1**, 127–134 (2005).
2. Vallet-Courbin, A. *et al.* A Recombinant Human Anti-Platelet scFv Antibody Produced in *Pichia pastoris* for Atheroma Targeting. *PLOS ONE* **12**, e0170305 (2017).
3. Hagemeyer, C. E. *et al.* Particle generation, functionalization and sortase A-mediated modification with targeting of single-chain antibodies for diagnostic and therapeutic use. *Nat. Protoc.* **10**, 90–105 (2015).
4. Oumzil, K. *et al.* Solid Lipid Nanoparticles for Image-Guided Therapy of Atherosclerosis. *Bioconjug. Chem.* **27**, 569–575 (2016).

ANNEXES

RESEARCH ARTICLE

A Recombinant Human Anti-Platelet scFv Antibody Produced in *Pichia pastoris* for Atheroma Targeting

Amelie Vallet-Courbin¹, Mélusine Larivière², Agnès Hocquet¹, Audrey Hemadou², Sarjapura-Nagaraja Parimala¹, Jeanny Laroche-Traineau², Xavier Santarelli¹, Gisèle Clofent-Sanchez², Marie-Josée Jacobin-Valat^{2‡}, Abdelmajid Noubhani^{1‡*}

1 UMR5248, CBMN, Bordeaux-INP, Pessac, France, **2** Centre de Résonance Magnétique de Systèmes Biologiques, Centre Nationale de Recherche Scientifique et Université de Bordeaux, Bordeaux, France

‡ These authors are co-last authors on this work.

* noubhani@bordeaux-inp.fr



OPEN ACCESS

Citation: Vallet-Courbin A, Larivière M, Hocquet A, Hemadou A, Parimala S-N, Laroche-Traineau J, et al. (2017) A Recombinant Human Anti-Platelet scFv Antibody Produced in *Pichia pastoris* for Atheroma Targeting. PLoS ONE 12(1): e0170305. doi:10.1371/journal.pone.0170305

Editor: Pablo Garcia de Frutos, Institut d'Investigacions Biomediques de Barcelona, SPAIN

Received: July 11, 2016

Accepted: January 3, 2017

Published: January 26, 2017

Copyright: © 2017 Vallet-Courbin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper.

Funding: This study was funded by (1) the French National Research Agency Program ANR-07-PCVI-0023 IMATHABIO, (2) the SFR TecSan (Technology for Health) of University of Bordeaux and (3) a public grant from the French National Research Agency within the context of the Investments for the Future Program, referenced ANR-10-LABX-57 and named TRAIL, Grant ANR-

Abstract

Cells of the innate and adaptive immune system are key factors in the progression of atherosclerotic plaque, leading to plaque instability and rupture, potentially resulting in acute atherothrombotic events such as coronary artery disease, cerebrovascular disease and peripheral arterial disease. Here, we describe the cloning, expression, purification, and immunoreactivity assessment of a recombinant single-chain variable fragment (scFv) derived from a human anti- α IIb β 3 antibody (HuAb) selected to target atheromatous lesions for the presence of platelets. Indeed, platelets within atheroma plaques have been shown to play a role in inflammation, in platelet-leucocyte aggregates and in thrombi formation and might thus be considered relevant biomarkers of atherosclerotic progression. The DNA sequence that encodes the anti- α IIb β 3 TEG4 scFv previously obtained from a phage-display selection on activated platelets, was inserted into the eukaryote vector (pPICZ α A) in fusion with a tag sequence encoding 2 cysteines useable for specific probes grafting experiments. The recombinant protein was expressed at high yields in *Pichia pastoris* (30 mg/L culture). The advantage of *P. pastoris* as an expression system is the production and secretion of recombinant proteins in the supernatant, ruling out the difficulties encountered when scFv are produced in the cytoplasm of bacteria (low yield, low solubility and reduced affinity). The improved conditions allowed for the recovery of highly purified and biologically active scFv fragments ready to be grafted in a site-directed way to nanoparticles for the imaging of atherosclerotic plaques involving inflammatory processes and thus at high risk of instability.

Introduction

Atherosclerosis is an inflammatory disease associated with the formation of unstable thrombosis-prone atheroma plaques made of large lipid cores, thin fibrous cap and inflammatory cell infiltrates within the walls of arteries.[1] Atherosclerotic plaque rupture is the mechanistic cause of about 75% of all sudden and often fatal heart attacks.[2] As the risk of plaque rupture

13-BSV5-0018 French National Research Agency Program.

Competing Interests: Pall Life Sciences provided support on the Octet Red experiments. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

Abbreviations: A-PL, Activated Platelets; BLI, Bio-Layer Interferometry; ELISA, Enzyme-Linked Immunosorbent Assay; HuAb, human antibody; IHC, Immunohistochemistry; IMAC, Immobilized Metal Affinity Chromatography; NA-PL, Non Activated Platelets; OD, Optical Density; PRP, Platelet Rich Plasma; RU, Relative Units; scFv, single-chain variable fragment; SPR, Surface Plasmon Resonance; VH, Heavy Chain; VL, Light Chain.

is more related to the plaque contents than to the plaque size, molecular imaging modalities have risen as a new imperative. Current studies tend towards the development of non-invasive targeted methods to assess the cellular components that underlie the risk of rupture.[3,4] Molecular imaging requires highly sensitive and specific probes made of a signal detection compound and an affinity ligand for targeting. The affinity ligand should be able to recognize molecules and cells over-expressed during the course of atherogenesis. Inflammation is a well-recognized pathophysiological process involving both innate and adaptive immune cells.[5] Recruitment of monocytes in the vascular wall and macrophage differentiation and proliferation represent a hallmark in the pathology of atherosclerotic lesions.[6] They contribute to the processes that underlie atherogenesis such as lipid accumulation, secretion of pro-inflammatory cytokines, extracellular matrix remodelling. Moreover, the observation of activation and oligoclonal expansion of T cells has suggested the presence of inciting antigens (Ags) that sustain T cell recruitment within coronary lesions.[7] B cells also play a pro or anti-atherogenic role depending on the subtypes (B1(a) or B2), and in atherosclerosis they accumulate both in the atherosclerotic intima and associated adventitia.[8–10] More recently, platelets have come to the forefront as partners of macrophages, T cells and B cells in inflammation and immune responses. They are now recognized as key players in innate and adaptive immune responses [11,12] and notably shown to modulate the T-effector/T-regulator balance via the CD40 ligand.[13,14] Platelet-derived CD40 ligand has also been reported to support B-cell differentiation and immunoglobulin class switching in mice.[15] Several cytokines released by activated platelets have been demonstrated to modulate monocyte and macrophage function.[16] Moreover platelet—leukocyte interactions largely contribute to OxLDL uptake and foam cell formation.[17] A recent study has underlined the presence of platelets not only in thrombi and intraplaque hemorrhage but also in atheroma burden, around necrotic areas and neovessels, shedding light on the rationale for targeting platelets within atherosclerotic lesions.[18]

Today, antibodies are used for several applications in research, diagnostics, and therapy. [19] Technology improvements are focused on several approaches to manufacturing recombinant human antibodies.[20] Moreover, *in vitro* selection technologies such as antibody phage display or ribosomal display have accelerated the generation of these recombinant human antibodies.[21–23]

To develop a novel non-invasive targeting approach for atheroma, our team previously selected, using *in vitro* phage display biotechnology on activated platelets, a phage-scFv fully human antibody (HuAb) specific to the α IIB β 3 integrin, which is an integrin only expressed on platelets and not on other immune cells.[24] This human antibody was further processed as a whole human IgG₄ molecule in baculovirus system.[18] We proved the maintenance of the bioactivity after grafting onto superparamagnetic nanoparticles dedicated to MRI imaging. However, the chemical functionalization was hard to proceed, time-consuming and we did not succeed in grafting more than one HuAb onto each nanoparticle [18]. To overcome these drawbacks and obtain a better conjugation ratio, another type of protein engineering has been applied to reduce the probe size. A scFv protein composed of the heavy (VH) and light (VL) chains of an antibody linked with a flexible peptide, has been constructed by recombinant DNA technology. The diameter of scFv fragments (5 nm), one-fifth the size of whole IgG, is more suitable for functionalizing relatively small nanoparticles. Compared to much larger forms of antibodies such as Fab, Fab'2 and IgG, scFv have lower retention times in non-target tissues and exhibit more rapid blood clearance and better penetration into targeted lesions. [25] Moreover, this recombinant form can be generated with tags for purification and site-specific attachment via engineered thiols to avoid loss of bioactivity. There are a variety of recombinant production systems for the generation of scFvs ranging from bacteria, filamentous fungi, insect cell lines to transgenic plants.[20,26] More specifically, studies reported the

production of an anti-platelets scFv antibody in transient mammalian cells system (e.g, free-style HEK 293F cells).[26] However, transient antibody production appears more suitable for small-scale production in antibody screening. Here, TEG4 HuAb needs to be stably expressed as scFv fragments in a quantity so that purification, characterization and further grafting of isolated recombinant fragments could be readily accomplished. Hence, TEG4 HuAb was processed in *Pichia pastoris*. The choice of production in yeast cells was guided by previous experiments where biological activity could only be recovered in the cytoplasm of cells when using *E. coli* as a host expression system. Indeed, despite improvements in fedbatch fermentation production, a proportion of inactive scFv remained in the total purified population.[27] From this point of view, the *Pichia pastoris* expression system is an attractive way due to its ability to secrete a large amount of recombinant protein in the supernatant, thereby facilitating the purification steps.

The aim of the study was thus to produce highly purified and biologically active scFv fragments in a suitable yeast cell expression modality. The biochemical characteristics of the anti- α IIB β 3 scFv were evaluated by ELISA, affinity binding analyses, flow cytometry against platelets and immunohistochemistry (IHC) on atheroma plaques from animal models and human coronary sections. These scFv fragments are ready to be grafted in a site-directed way to nanoparticles for the imaging of atherosclerotic plaques where inflammatory and immune processes increase the risk of instability.

Materials and Methods

Materials

Large-scale expression of TEG4-2c scFv in *Pichia pastoris* was carried out in BIOSTAT Bplus laboratory fermentor 5L (Sartorius-Stedim Biotech, Germany). Data and set points were monitored with MFCS-Win. Growth media were purchased from Becton Dickinson (Le Pont de Claix, France).

Immobilized Metal Affinity Chromatography (IMAC) was carried out on HisTrap™ excel column (id 5 mL, GE Healthcare, Sweden) containing resin charged with nickel ions. Buffers for chromatographic runs, and reagents were prepared using chemicals of analytical grade from Sigma-Aldrich (Saint Quentin Fallavier, France).

Chromatographic experiments were performed using ÄKTA pilot workstation (GE Healthcare) and were monitored with Unicorn 5.1 software. Protein detection was monitored at 215 and 280 nm.

Spectrophotometric measurements of samples drawn from the fermentor and BCA protein assay measurements were carried out using SAFAS UVmc2 double-beam UV-visible spectrophotometer (Société Anonyme de Fabrication d'Appareillages Scientifiques, Monaco). Flow cytometry experiments were performed on a FACSCanto I cytometer from BD Biosciences (Le Pont de Claix, France) and monitored using DIVA software.

The absorbance of ELISA assays was read at 405 nm using a CHAMELEON microplate reader from ScienceTec (Les Ulis, France).

Surface Plasmon Resonance (SPR) experiments on purified α IIB β 3 integrin (Kordia Life Sciences, Leiden, The Netherlands) were carried out using the BIAcore™ 3000 (GE Healthcare Europe GmbH, Velizy-Villacoublay, France) equipped with research-grade CM5 sensor chips. Bio-Layer Interferometry (BLI) experiments on lyophilized platelets (Helena Biosciences Europe, Queensway S, UK) were performed using an Octet instrument (Octet Red96 Pall Life Sciences, Saint-Germain-en-Laye, France) and HIS2 (anti-penta Histidine Ab) sensors.

Methods

1-Vectors and Strains. The pHOG21 vector, kindly provided by M. Little (Affimed Therapeutics, Ladenburg, Germany), pCR4-TOPO and pPICZαA (Invitrogen, Carlsbad, USA) plasmids were used for preparation of DNA constructs. Synthetic oligonucleotides were purchased at Eurogentec (Liege, Belgium). Cloning step was carried out with *Escherichia coli* strain JM109 (NewEngland Biolabs, Ipswich, MA USA). *Pichia pastoris* strain X33 used for production of recombinant TEG4-2c scFv was obtained from Invitrogen (Carlsbad, CA, USA).

2-Construction of an expression vector for TEG4-2c scFv. The TEG4 scFv DNA fragment provided in the pHOG21 vector previously used for the TEG4 scFv expression in XL1-Blue *E. coli* [16] was used as template to generate a new scFv format containing 2 cysteines into the C-terminal end. The coding sequence was PCR-amplified with *Pfu* Turbo polymerase (Stratagene). The primers (P1CTE: 5' -TATCACGTGGCAGGTGCAGCTGGTGG-3' and P1C2C: 5' -TCTAGATTAGCAGCACCCGTGATGGTGGTGG-3') were used to introduce the 6HisTag- Gly-Cys-Cys-Stop amino acid sequence. Thirty amplification cycles were performed (30s at 95°C, 45s at 48°C and 1min at 72°C) followed by a final extension of 10 min at 72°C. The PCR product was purified, ligated into pCR4-TOPO plasmid and verified by DNA sequencing (Millegen Technology—France). The resultant plasmid was digested by *Pml*I and *Xba*I to excise TEG4-6His-Gly-Cys-Cys DNA fragment and ligated into the expression vector pPICZαA (Invitrogen, Carlsbad, USA) containing the Zeocin resistance gene for selection and the AOX1 promoter. The new construction of TEG4-2c scFv is shown in Fig 1.

The resultant plasmid was linearized by *Pme*I and transformed into competent *P. pastoris* X-33 cells by electroporation using Gene Pulser II (Bio-Rad, Hercules, CA, USA). Transformant cells were grown on YPDS-Zeocin 150 µg/mL-agar plates and screened again later for their ability to grow on Zeocin (Invitrogen, Carlsbad, USA) up to concentrations of 2 mg/ml.

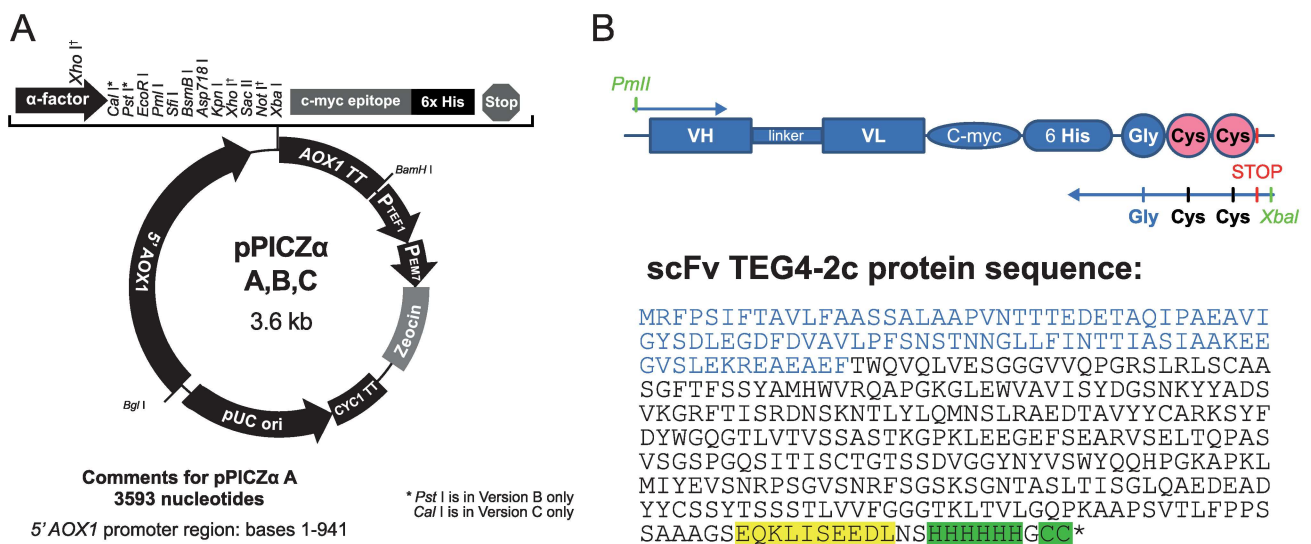


Fig 1. pPICZαA expression vector according to the EasySelect *Pichia* Expression Kit Manual (Invitrogen) and a schematic representation of TEG4-2c scFv with the protein sequence. (A): All the featured restriction sites are unique. 5' AOX1: promoter region of AOX1; TT AOX1: transcription termination of AOX1; PTEF1: promoter of TEF1; PEM7: promoter of EM7; Zeocin resistance: Sh ble ORF; CYC1 TT: transcription termination of CYC1. **(B):** The TEG4-2c scFv coding sequence was cloned between *Pml*I and *Xba*I sites and the protein sequence of recombinant tag-scFv including the 6His-tag and the 2 cysteine are highlighted in green. The α-factor signal sequence is represented in blue and the C-myc tag is highlighted in yellow.

doi:10.1371/journal.pone.0170305.g001

Selected clones were tested for the presence of TEG4-2c scFv coding sequence in their genome by colony PCR analysis using the PICTE and PIC2C primers, amplification cycles were preceded by a heating step 3 min at 95°C.

3-Expression and purification of TEG4-2c scFv. Transformed *P. pastoris* cells that exhibited high resistance to Zeocin (up to 2 mg/mL plates) were grown in shake flasks containing 100 ml of buffered glycerol complex medium (BMGY, 1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer at pH 6.0, 13.4g/L YNB, 4×10^{-4} g/L biotin, 10 g/L glycerol and 150 µg/mL Zeocin) until an optical density of 20 was reached. The culture was inoculated in a 5L bioreactor at 0.2 OD units. The bioreactor conditions were optimized by modification of Narasimhan, J *et al.*[28] The temperature and pH were maintained at 30°C and 6 respectively, and dissolved oxygen levels were maintained at 20% saturation by regulating aeration and agitation in a cascading system. After complete consumption of glycerol in the medium (20 to 24 hours), a methanol fed-batch phase was initiated by adding methanol every 12 h to a final concentration of 0.6%; this phase promotes the induction of scFv production and secretion in the medium. The methanol feed frequency was modified to every 6 h during the 4 latest days of production. Samples were drawn every 24 h to determine the yeast growth profile and scFv production. After 120 h of induction, the culture was harvested and the supernatant frozen at -80°C pending purification steps. Prior to purification, the pH of the broth was adjusted to 7.4 and filtered through a 0.45 µm cellulose acetate membrane.

The expressed recombinant TEG4-2c scFv was purified using IMAC. The HisTrap Excel resin was equilibrated with 50 mM Tris-HCl pH 7.5, 500 mM NaCl (buffer A at a flow rate of 3 mL/min). Typically, 450 to 800 mL *P. pastoris* expression broth supernatant containing the scFv was directly injected into the column. The column was then washed with buffer A until absorbance at 280nm reached the baseline. The elution was carried out into two steps using 5% and 30% buffer B (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 500 mM imidazole) corresponding respectively to 25 mM and 150 mM imidazole. The elution fraction was dialyzed against PBS (NaCl 135 mM, KCl 2.5 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 1.5 mM pH 7.4) buffer.

4-Protein assay, SDS-PAGE and Dot-blot analysis. Protein quantification of the chromatographic fractions was performed using bicinchoninic acid protein micro-assay (BCA kit, Sigma). Twenty-five microliters samples were incubated with 200 µL of BCA working reagent and plate was incubated at 37°C for 30 min. The absorbance was measured at 562 nm.

Fractions obtained from chromatographic experiments were analyzed by SDS-PAGE under reducing conditions over 12% polyacrylamide gels. Supernatant and flow through fractions were 5x concentrated by TCA precipitation. Precision plus protein prestained standards (BIO-RAD) were used as molecular weight ladder.

To analyse the efficiency of the yeast cells *P. pastoris* to produce the scFv fragments into the broth medium, 50 µL samples from day 1 to day 5 were blotted on a nitrocellulose membrane using Bio-Dot Microfiltration Instrument (BIO-RAD). The membrane was blocked with a blocking buffer (TBS Tween with 3% milk powder) for 2 hours. The membrane was again washed twice with TBS-tween, and then incubated with primary antibody (Anti 6His, SIGMA) at 1:1500 dilution overnight at 4°C. Membranes were then washed and probed with a secondary antibody (anti-mouse IgG-HRP Cell Signaling Technology) at 1:5000 dilution.

Colorimetric analysis was performed using Opti4CN (BIO-RAD) kit by gently shaking until color develops.

5-Preparation of Platelet-Rich-Plasma (PRP) and washed platelets. Platelet-rich plasma (PRP) and washed platelets were prepared from the blood of voluntary healthy donors.

PRP was obtained from venous blood anticoagulated with sodium citrate (0.38%, w/v) after centrifugation at 120 g for 10 min at room temperature.[29]

For the preparation of washed platelets, venous blood was anticoagulated with citric acid-citrate-dextrose NIH formula A (ACD-A) (1 vol of anticoagulant: 6 vol of blood). After centrifugation at 120 *g* for 10 min, the PRP was collected and mixed with ACD-A (1 vol: 9 vol, PRP) plus 100 nM prostaglandin E₁ and 0.05 U/mL Apyrase^{Grade7}, [30]. Platelets were sedimented by centrifugation at 1100 *g* for 15 min, washed and adjusted at 10⁸/mL in a modified Tyrode's buffer (137 mM NaCl, 2.7 mM KCl, 12 mM NaHCO₃, 0.3 mM NaH₂PO₄, 1 mM MgCl₂, 5.5 mM glucose, 5 mM HEPES, 0.1% (wt/vol) bovine serum albumin (BSA), pH 7.4).

6-Analysis of TEG4-2c scFv reactivity by Flow cytometry.

Analysis on washed platelets

For the analysis of TEG4-2c scFv reactivity on platelets, washed platelets were activated with 0.5 U/ml human α -thrombin (Fibrindex, Ortho-Diagnostics, Raritan, NJ). Samples of both activated and non-activated platelets were fixed for 30 min with an equal volume of paraformaldehyde (PFA) 2% and then incubated with diluted TEG4-2c scFv.

Aliquots of 10⁸ PFA-fixed-washed-platelets/mL (10 μ L) [27] non-activated (NA-PL) or activated with thrombin (A-PL) were incubated with 25 μ g/mL of TEG4-2c scFv human antibody or PAC-1 commercial IgM antibody (BD Biosciences) [31] targeting the activated α IIb β 3 integrin overnight at 4 °C. After two washes in PBS, a 30 min incubation with secondary Alexa Fluor 488 anti-6His or goat anti-mouse IgM antibodies (1:20) was performed for detection of respectively, the TEG4-2c scFv and the murine PAC-1 antibody. Reactions in the absence of antibodies were used as negative controls.

Analysis on PRP samples

Aliquots of 4 μ L of PRP are used in each experiment. Some aliquots are activated with 50 μ M thrombin receptor activating peptide (TRAP6; Sigma Aldrich, Saint-Quentin-Fallavier, France). Activated (+TRAP) and non-activated PRP aliquots (-TRAP) were then incubated with diluted TEG4-2c scFv (25 μ g/ml) 15 min before adding secondary Alexa Fluor 488 anti-6His (both 1:20) antibody for another 15 min incubation. Reactions in the absence of antibodies were used as negative controls.

Cells were then resuspended in PBS before analysis on the FACSCanto I cytometer. The forward and wide-angle light scattering and fluorescence intensity from 10,000 platelets were collected using a logarithmic gain.

7-Analysis of TEG4-2c scFv reactivity by ELISA. A 96-well flat-bottom polystyrene microtiter plate (Costar, Corning, NY) was coated with washed thrombin-activated (A-PL) or non-activated platelets (NA-PL) at 10⁷ platelets/well or BSA protein 20 μ g/mL (Sigma-Aldrich) in 50 mM carbonate-bicarbonate buffer (pH 9.6) overnight at 4 °C. The plate was washed three times with PBS (pH 7.4) containing 0.05% Tween 20 (200 μ L/well) and blocked with 5% skimmed milk in PBS for 2 h at 37 °C. TEG4-2c scFv fragments purified by IMAC or full-length mouse IgG antibody (AP2 antibody targeting α IIb β 3 integrin) [32] were tested on A-PL and NA-PL. After incubation for 2 h at 20 °C, the plate was washed and incubated with an anti-6His IgG antibody (GE Healthcare) for 1 h at 37 °C to detect scFv fragments or with PBS for full-length mouse antibodies. After washings, 100 μ L of a 1:1000 dilution of horseradish peroxidase (HRP)-conjugated anti-mouse IgG (Immunotech, Marseille, France) was added and incubated for 90 min at room temperature. Color was developed with 100 μ L of 2,2'-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) (Sigma-Aldrich, Saint-Quentin, Fallavier, France) and the absorbance was read at 405 nm using a CHAMELEON microplate reader.

8-Preparation of atheromatous and healthy aorta proteins. All animal experiments were performed in conformity with the Guide for the Care and Use of Laboratory Animals (NIH Publication No. 85–23, revised 1996) and were accredited by the local ethical committee.

Adult male New Zealand rabbits (NZW), weighting from 2.5 kg to 3.0 kg, were obtained from Charles Rivers Laboratories (St Germain sur l'Arbresle, France).

In order to mimic atherosclerosis that develop in humans, rabbits were fed a fat atherogenic diet including 0.3% cholesterol for 8 months and were subjected to surgeries to allow the formation of complex plaques with intramural thrombi. A first surgery was performed to remove endothelial cells from the thoracic until the abdominal aorta using a Fogarty catheter (Fogarty 4F; Edwards Lifesciences). The second surgery consists in an angioplasty using an expandable latex balloon (Maxxum, Boston Scientific; 20 mm long, diameter of 4.5 mm) under radioscopic guidance from the region of renal arteries to iliac bifurcation. Surgeries were performed under anesthesia by intramuscular administration of 20 mg/kg ketamine and 2 mg/kg xylazine. Anesthesia was maintained with isoflurane gas (0.25% to 0.35%). As a preventive anti-thrombotic treatment, 1000 μ l of heparine (Héparine Choay, Sanofi Synthélabo) was infused. After surgery 100 mg aspirine (Aspégic injectable, Sanofi Synthélabo) was administered as analgesia. Aortas from untreated rabbits and balloon-injured aortas from hypercholesterolemic rabbits were extracted from the aortic arch to the iliac bifurcation, washed and fractioned in order to solubilize tissue proteins in T-PER (Thermo Fisher Scientific) or HSB (50 mM HEPES, pH 7.4; 137 mM NaCl; 1% NP-40 (v/v); 2 mM EDTA; 1 mM PMSF; protease inhibitors cocktail (Roche Diagnostics, Meylan, France) lysis buffer. Homogenization was performed using first a Polytron TP-20 Homogenizer (Kinematica, Lucerne, Switzerland) and then a sonicator (3 x 10-seconds pulses at 80% magnitude). After two centrifugations at 13000 g for 45 min at 4°C to remove insoluble material from the supernatant, the protein concentration of every soluble extract was determined using the Bradford assay kit according to the manufacturer's instructions (Thermo Fisher Scientific).

9-Affinity determination evaluated by SPR. The interactions between the anti- α IIB β 3 TEG4-2c scFv and the integrin α IIB β 3 were analyzed by SPR sensing using Biacore™ 3000 (GE Healthcare), according to the manufacturer's instructions. HBS-EP buffer (0.01 M HEPES pH 7.4, 0.15 M NaCl, 3 mM EDTA, 0.005% v/v Surfactant P20) (GE healthcare life sciences, France) with 2 mM Ca²⁺ was used as a running buffer. The purified α IIB β 3 was first dialysed in the running buffer, diluted in 10 mM Na-Acetate, pH 5.5, then immobilized onto the sensor surface of a Biacore sensor chip CM5 to an immobilized ligand density (RU) of about 8000, using an amine coupling kit. Then, anti- α IIB β 3 TEG4-2c scFv diluted into the running buffer at concentrations ranging from 90 to 400 nM was allowed to flow through the sensor chip for 5 min at a flow rate of 20 μ l/min and dissociation of bound analyte was allowed to proceed for 15 min before chip regeneration with NaOH 20mM.

10-Affinity determination evaluated by BLI: Octet Red96. Octet instrument is a label-free detection system that exploits optical principle to read bimolecular interactions, the bio-layer interferometry (BLI). The interactions between the anti- α IIB β 3 TEG4-2c scFv and blood platelets were analyzed using Octet (Red96) in platelet buffer (NaCl 137 mM, KCl 2 mM, NaH₂PO₄ 0.3 mM, MgCl₂ 1 mM, Glucose 5.5 mM, HEPES 5 mM, Bicarbonate-Na 12 mM pH6). The purified TEG4-2c scFv was loaded on a HIS2 biosensor (anti-penta Histidine Ab optical fiber based sensor) at 21 μ g/mL. Platelets were diluted in PBS, coated into 96 well plate and analyzed at different concentrations: 5x10⁵ to 5x10⁸/mL, corresponding to integrin α IIB β 3 molarities from 42 pM to 42 nM. The plate is shaken during reading to create an "orbital flow". Controls with no scFv and no platelets were performed to check for nonspecific binding and signal drift, respectively.

11-Immunohistochemistry Analysis on Murine and Human Atherosclerotic Sections. All animal studies were approved under the N°50120192-A by the Animal Care and Use Committee of Bordeaux, France. The studies were performed in accordance with the ethical standards laid down in the declaration of Helsinki. ApoE^{-/-} mice were fed a high-cholesterol diet for 21 weeks to allow for the development of atherosclerotic lesions, afterward they were terminally anaesthetized using a lethal dose of isoflurane. The chest was opened and the aorta

retrieved from the aortic arch to the iliac bifurcation. The lesions-rich areas were subsequently isolated, fixed and embedded in paraffin. New Zealand white rabbits were prepared as described in Methods, section 8.

Human coronary arteries were harvested from patients with end-stage heart failure having undergone heart transplantation. All of the clinical interventions took place at Haut-Lévêque Hospital (Pessac, France). All work with tissues from human subjects had been approved by the CPP committee (Comité de Protection des Personnes Sud-Ouest et Outre Mer) of Bordeaux and from the Research Ministry in France (Authorization number DC -2016- 2724). The CPP committee waived the need for patient written consent because surgical waste no longer attached to the person is considered "RES NULLIUS". Patients were informed by the clinicians; if they did not express their opposition to research, the de-identified samples were immediately processed and embedded in paraffin.

Paraffin-embedded sections of arterial tissue from mouse, rabbit or human were used in IHC experiments. The following antibodies were tested on sections from each specie: TEG4-2c scFv (human fragment antibody, tested between 20 and 40 µg/ml), anti-CD41 (anti-αIIb rabbit antibody, working dilution 1:200, Abcam, France), AP2 (anti-αIIbβ3 mouse antibody, 10 µg/mL), RAM11 (mouse antibody, working dilution 1:50) and PGM1 (mouse antibody, working dilution 10µg/ml) anti-CD68 antibodies targeting rabbit and human macrophages respectively (Dako, Les Ulis, France).

Paraffin-embedded sections were deparaffinised, rehydrated, and heat mediated antigen retrieval was performed with Tris-EDTA pH 9 buffer following the specifications of Abcam, Paris, France (www.abcam.com/ps/pdf/protocols/ihc_p.pdf). Endogenous peroxidase was then blocked with 3% H₂O₂ in water, for 15 min. After washing in PBS 1X + 0.025% Triton (PBST), nonspecific binding was blocked with PBS 1X + 0.2% Triton + 2% bovine serum albumin (BSA) for 1 h at room temperature.

Afterwards, antibodies were applied overnight at 4°C, diluted at 20 µg/mL in PBS 1X + 1% BSA. The following day, three washes with PBST were performed. To detect TEG4-2c scFv fragments, an HRP-conjugated antibody specific to 6His (working dilutions 1:250) was applied to the sections for 1 h at room temperature. HRP-conjugated secondary antibodies specific to (1) rabbit IgG H+L (Beckman Coulter), (2) mouse IgG H+L (Beckman Coulter) were respectively applied for commercial primary (1) anti-CD41, (2) RAM-11, PGM1 and AP2 antibodies.

After a further three washes with PBST, staining was performed by adding the peroxidase substrate diaminobenzidine (DAB substrate kit, Eurobio/ABCys, Les Ulis, France) with H₂O₂. It yielded a yellow brown deposit within 10 min at room temperature. After a wash in dH₂O to stop the enzymatic reaction, slides were counterstained in hematoxylin, dehydrated and mounted.

Results

1-Monitoring TEG4-2c scFv large scale production in *Pichia pastoris*

After the transformation of *P. pastoris* with linear pPICZαA-TEG4-2c plasmid, up to 60 clones were selected on 2 mg/mL Zeocin agar plates. The colony PCR analysis confirmed the presence of TEG4-2c scFv gene in 16 tested clones. Six clones were evaluated for their capacity to produce and secrete TEG4-2c scFv in a shake flask. Finally, the clone TEG4-2c A2 was selected for further analysis and the production scaled up in a 5 L benchtop bioreactor.

Five culture batches were carried out on a 5 L benchtop bioreactor and exhibited a similar growth profile (data not shown). The culture conditions were maintained as described in the experimental section. The dissolved oxygen level was maintained throughout the production

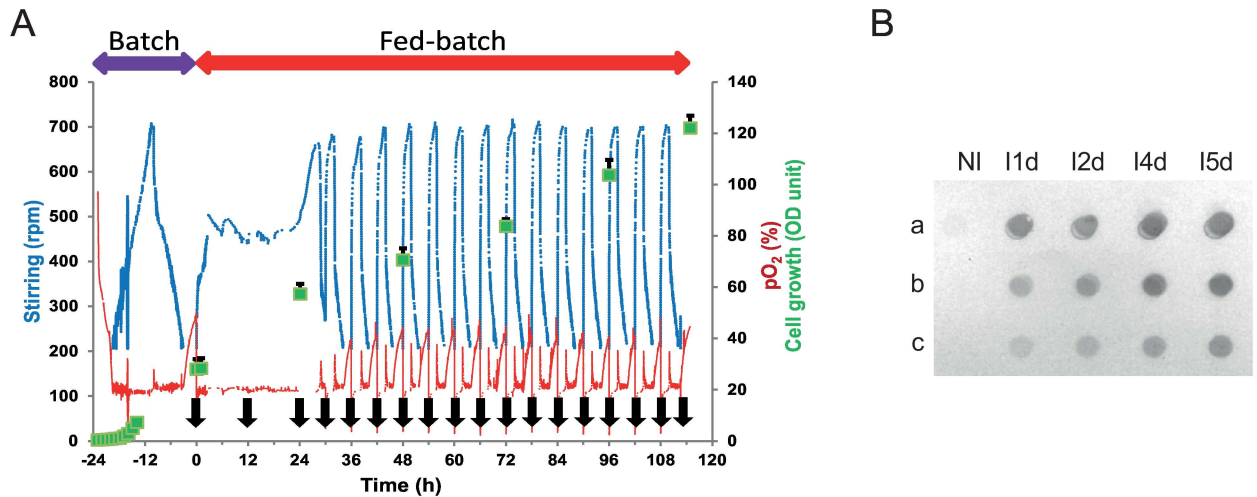


Fig 2. TEG4-2c scFv production process. A: Fed batch fermentation history plot. Stirring, pO_2 and OD_{600} values are plotted versus time during the cultivation of *P. pastoris* in BMGY medium. Cultures were induced with methanol at $t = 0$ (24 h after starting the batch phase) during the fed batch phase the methanol was added every 12 h or 6h (black arrows) to a final concentration of 0.6%. The average values are shown with error bars representing the standard deviation ($n = 5$). 1 OD_{600} unit was equivalent to 0.29 mg/mL dry weight. **B: Dot-blot analysis of supernatants from recombinant *P. pastoris* culture.** Fifty microliters samples from non-induced culture (NI) and from day 1 to day 5 induced cultures (I1d to I5d) were undiluted (a) or diluted (b = 1:10; c = 1:50) and blotted on a nitrocellulose membrane. The recombinant TEG4-2c scFv were detected with the Anti-6His antibody and revealed by colorimetric analysis.

doi:10.1371/journal.pone.0170305.g002

batch at 20% by a combination of agitation and aeration systems, in a cascading manner to improve the oxygen supply during the growth of cells (Fig 2A). The yeast growth was exponential during the batch culture with glycerol as the substrate; the glycerol was exhausted after 20 to 24 h culture and the cell density reached 28 OD units corresponding to 8.1 g/L (dry weight).

A fed-batch phase was initiated by adding methanol every 12 h to induce the scFv production. To optimize the efficiency of scFv production during the 4 last days, the methanol feed frequency was every 6 h. Following each substrate addition event, there was a sharp increase in agitation indicating the active consumption of substrate and sustainable growth of cells. Stirring, which again diminished upon substrate consumption, was raised following the consecutive methanol injection. During the fed-batch phase, a linear increase in yeast biomass was observed 24 h after induction. A growth rate of $0.23 \text{ g L}^{-1} \text{ h}^{-1}$ was maintained during 5 days; the cell density was also increased by a 4.3 fold and reached 35 g/L (dry weight).

Dot-blot analysis was performed to evaluate the TEG4-2c scFv expression each day after induction in the culture broth. Recombinant TEG4-2c was only produced upon induction of transformed X-33 cells with methanol. This was clearly confirmed by the absence of signal into the samples before methanol feeding and into the samples from X33 *Pichia* cells transfected with the empty pPICZ α A plasmid. The anti- α Iib β 3 TEG4-2c scFv was expressed from the first day in a soluble form into the cells culture medium (Fig 2B).

2-Purification of TEG4-2c scFv by IMAC

For the purification of recombinant TEG4-2c scFv, optimal performance was obtained when an intermediate washing step with 25 mM imidazole was used to remove the contaminating proteins weakly bound to the column, after which recombinant TEG4-2c scFv was eluted using 150 mM imidazole (Fig 3A).

Typically, 30 mg of TEG4-2c scFv were produced and secreted by *P. pastoris* into 1L broth medium after 5 days of growth. The scFv content into this supernatant was estimated to 1.4%

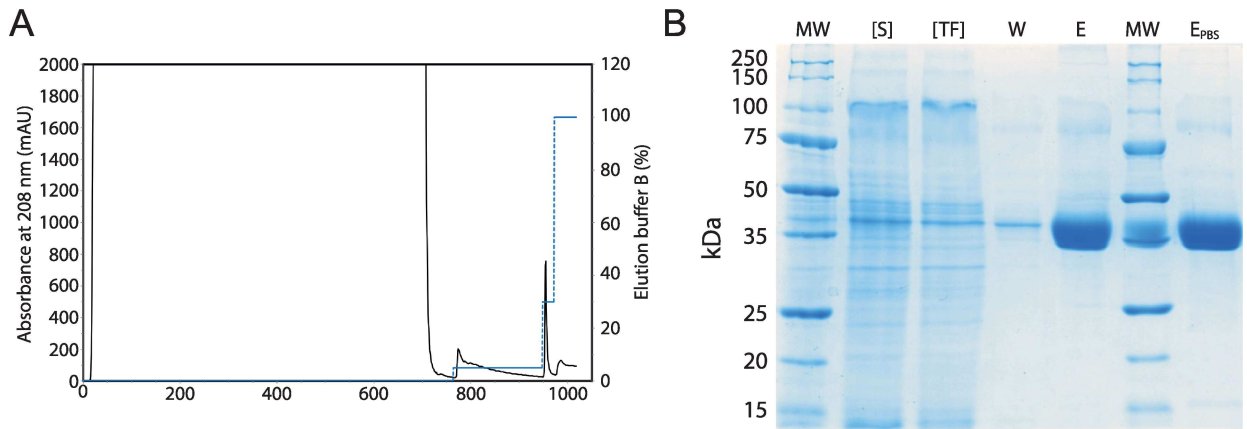


Fig 3. Purification of recombinant TEG4-2c scFv A: IMAC chromatogram. The HisTrap Excel resin (5 mL) was equilibrated with 50 mM Tris-HCl pH 7.5, 500 mM NaCl (buffer A at a flow rate of 3 mL/min). *Pichia pastoris* expression broth supernatant containing the TEG4-2c scFv was injected into the column. The column was then washed with buffer A until absorbance at 280nm reached the baseline. (Dot Line): The elution was carried out in two steps using 5% and 30% buffer B (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 500 mM imidazole) corresponding respectively to 25 mM and 150 mM imidazole. **B: Electrophoretic analysis of one step IMAC purification of recombinant TEG4-2c scFv.** 12% SDS-PAGE stained with colloidal blue, MW: molecular weight ladder (kDa). [S]: 5x concentrated culture supernatant. [TF]: 5x concentrated flow-through. W: 25 mM imidazole washing fraction. E: 150 mM elution fraction. E_{PBS}: Elution fraction dialyzed against PBS.

doi:10.1371/journal.pone.0170305.g003

of the total proteins. After the IMAC purification, the yield was around 22 mg TEG4-2c scFv from 1L culture medium. This represents a recovery of 70% of produced scFv with a high purity (> to 80%). This one step chromatography led to a 57 fold purer product with highly concentrated solution of TEG4-2c scFv (Table 1).

The eluted fractions were analyzed by SDS-PAGE (Fig 3B), which revealed a major band of 35 kDa from the elution step fractions, corresponding to the expected molecular mass of recombinant TEG4-2c scFv. A weaker band of 75 kDa was also present in this lane, but the mass spectroscopy analysis revealed that TEG4-2c scFv was the major protein present in this fraction (data not shown). This data strongly suggests that the higher molecular weight protein was indeed TEG4-2c scFv dimer.

The elution fraction was dialyzed against PBS buffer. The SDS-PAGE (Fig 3B-E_{PBS}) shows the purified TEG4-2c scFv as a single thick band with an estimated purity higher than 80%. The profile of TEG4-2c scFv into PBS buffer (used for ELISA and affinity measurement) is

Table 1. IMAC Purification of scFv TEG4-2c produced and secreted by *P. pastoris*. Data are standardized for 1 L culture media; the values are the mean of 7 independent experiments ± SD values.

	Protein ^a (mg/mL)	scFv ^b (µg/mL)	Volume (mL)	Total protein ^a (mg)	Total scFv ^{ab} (mg)	Step recovery (%)	Step purification (fold)
Supernatant	2.190 ± 0.367	30.7 ± 4.3	1000	2190 ± 367	30.7 ± 4.3	100	0
Flow through	1.983 ± 0.377	ND	1000	1983 ± 377	ND	ND	ND
Wash step	0.167 ± 0.036	16.7 ± 3.3	75 ± 5	12.53 ± 1.64	1.253 ± 0.25	ND	ND
Elution step	0.866 ± 0.136	695 ± 35	31 ± 2	26.85 ± 1.52	21.56 ± 3.01	70.3 ± 0.6	57.35

^a Protein concentration was determined by BCA protein assay, using bovine serum albumin as standard.

^{b,c} TEG4-2c scFv into the supernatant and the purity was estimated by densitometric quantification of corresponding lane of SDS-PAGE 12% acrylamide gel stained by colloidal blue, thanks to Bio-Rad Image Lab™.

ND: not determined.

doi:10.1371/journal.pone.0170305.t001

similar to the elution fraction (E) obtained at the end of IMAC. The final overall yield of TEG4-2c was 0.7 mg/mL; it represents 70% of initial product. However, upon SDS-PAGE analysis, it was observed that some amount of our protein of interest was also lost in the flow-through and washing steps.

3-Analysis of TEG4-2c scFv binding to platelets by flow cytometry and ELISA tests

3.1-Flow cytometry analysis using TEG4-2c scFv. Flow cytometry analysis showed a better recognition of washed activated platelets versus non-activated ones (Fig 4A). PAC-1, an IgM murine antibody specific to the activated conformation of α IIB β 3 was included as a positive control of the experiment. The slight recognition of resting platelets might be due to their activation during the purification process. Activation of platelets is a problem classically encountered during their handling and processing. To minimize manipulation steps, the binding specificity of TEG4-2c scFv was checked on platelets within plasma (analysis on PRP). The results (Fig 4B) highlighted a binding specificity for platelets activated in PRP with TRAP peptide whereas resting platelets were not recognized.

3.2-ELISA tests. The reactivity of TEG4-2c scFv produced in *Pichia pastoris* was then measured on human thrombin-activated and non-activated platelets and on purified α IIB β 3 by ELISA (Fig 4C). We have to underline that, here again, the coating of platelets on ELISA wells may itself induce their activation. However, a better recognition of activated platelets is reported with TEG4-2c scFv, especially when using lower concentrations. AP2, a murine antibody directed against α IIB β 3 was included as a positive control of the experiment.

4-Binding of scFv TEG4-2c to α IIB β 3 investigated by SPR and BLI

Affinity of TEG4-2c scFv was determined either by surface plasmon resonance (SPR) with a BIACORE device on purified α IIB β 3 or using the [Bio-Layer Interferometry](#) (BLI) technology with an OCTET instrument on lyophilized platelets resuspended from a freeze-dried solution.

4.1-BIACORE analysis. In this first experimental session using SPR technology, TEG4-2c scFv was injected at 3, 6 and 12,5 μ g/ml on immobilized α IIB β 3, corresponding respectively to 94, 188 and 390 nM. The binding was concentration-dependent (Fig 5A) but the three sensorgrams could not be fitted using the simple Langmuir model, likely because of the presence of a more complex interaction, implying monomers and dimers in the purified fractions.

4.2-Binding by BLI. This second set of experiments performed using the Octet instrument aimed at determining the interactions between TEG4-2c scFv and blood platelets (Fig 5B). This is more informative because this strategy allows evaluating the binding on α IIB β 3 in its natural conformation. TEG4-2c scFv immobilized through an anti-6His coating on an optical fiber based sensor was immersed in a solution of platelets (seven different concentrations were used) contained in the well of a 96-well plate. The plate is shaken during reading to create an "orbital flow". The different platelet concentrations, from 5×10^5 to 5×10^8 /ml, were converted into integrin α IIB β 3 molarities by taking into account the number of α IIB β 3 per platelet (50 000). Controls with no scFv allowed checking for non-specific binding of the platelets on the sensors. These controls were subtracted from the curves obtained with TEG4-2c scFv. Other controls were also performed with no platelets (only buffer), assessing the signal drift due to the potential release of scFv from sensors.

Octet experiments showed good interactions between TEG4-2c scFv and platelets at nanomolar concentrations superior to 0.4 nM. No Kd could be calculated because no dissociation was observed. However, considering the binding at 0.8 and 4 nM, the affinity should be in the nanomolar range.

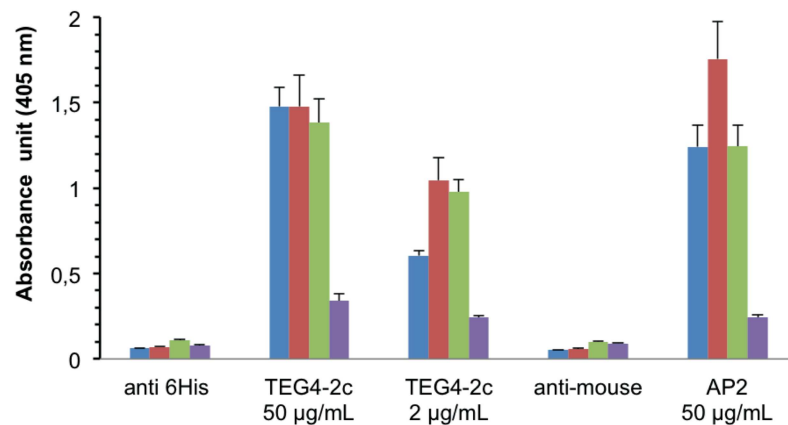
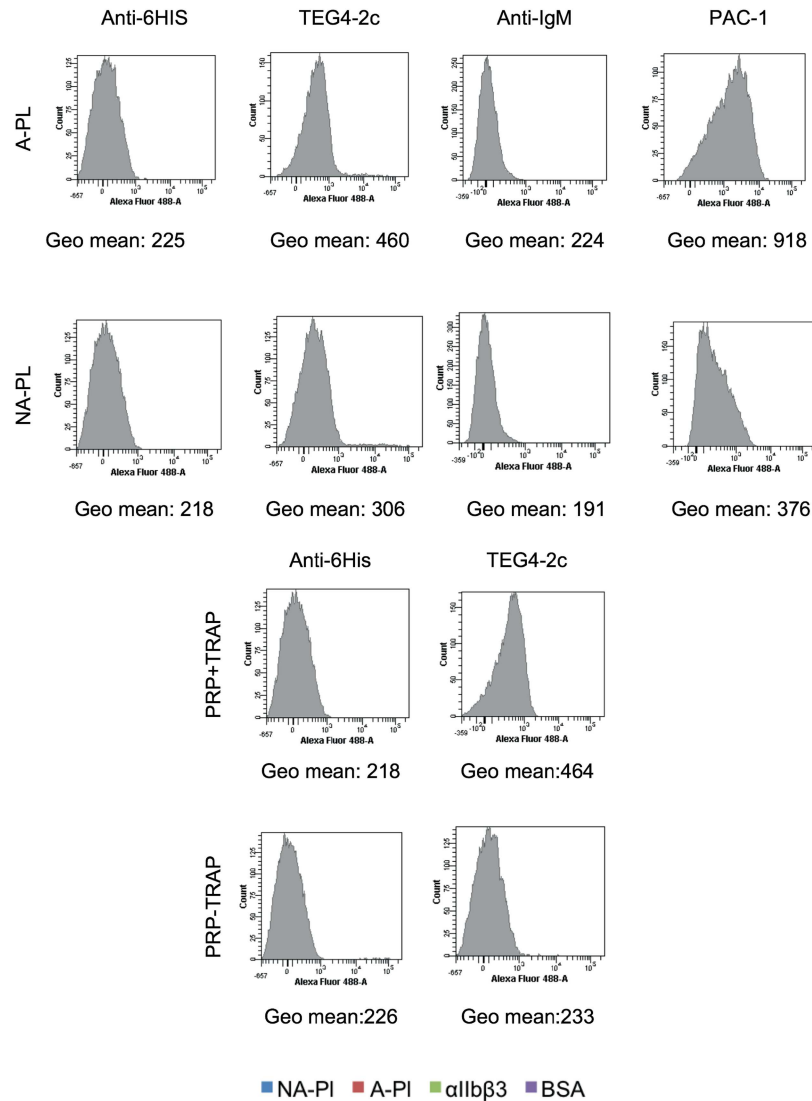


Fig 4. Binding assessment of TEG4-2c scFv to human platelets by flow cytometry and ELISA tests. A: Binding of TEG4-2c scFv on thrombin-activated human (A-PL) or non-activated—platelets (NA-PL) analysed by flow cytometry. PAC-1 IgM murine antibody serves as a positive control. Binding of antibody to the platelets was further detected by incubation with Alexa Fluor 488 anti-6His or anti-mouse IgM antibodies. Negative controls were secondary antibody only. Histograms depict representative data \pm SD of three

independent experiments. Quantitative fluorescence intensities (in Geo mean) are stated under each respective histogram. **B:** Binding of TEG4-2c scFv on TRAP-activated-human (+ TRAP) or non-activated platelets (-TRAP) analysed by flow cytometry. Quantitative fluorescence intensities (in Geo mean) are stated under each respective histogram. **C:** Representative whole cell (A-PL, NA-PL) and purified proteins (α IIb β 3, BSA) ELISA with TEG4-2c scFv. A murine anti- α IIb β 3 antibody AP2 was used as positive control. Negative controls were secondary antibody only. Binding of antibodies was visualized via HRP-6His or HRP-anti-mouse IgG. OD value represents absorbance at 450 nm. Plots represent the mean values \pm SD (n = 3)

doi:10.1371/journal.pone.0170305.g004

5-Evaluation of atheroma burden recognition by IHC and ELISA tests

As the final aim of the project is to target platelets colonizing the plaque, we evaluated the ability of TEG4-2c scFv to recognize platelets within aorta lesional sections of animal models of atherosclerosis as well as aorta sections recovered from human coronary specimens (Fig 6A). The reactivity was also tested on proteins isolated from the extracted aorta from atheromatous and healthy rabbits by ELISA (Fig 6B).

5.1-IHC. The immunohistochemistry data were in accordance with ELISA with a huge recognition of platelets in all analyzed sections (Fig 6A). An antibody directed against murine α IIb (anti-CD41) and antibodies directed against rabbit and human macrophages (anti-CD68) over-expressed in the atheroma burden were included as positive controls. We also compared the targeting of TEG4-2c scFv with that of AP2, an antibody we have already demonstrated able to label mouse, rabbit and human platelets.

5.2-ELISA tests. ELISA assay clearly showed a better recognition of atheromatous proteins with TEG4-2c scFv and AP2 antibodies (Fig 6B). An anti-CD68 macrophage rabbit antibody was included as positive control. We confirmed by mass spectroscopy analysis (data not shown) that the α IIb β 3 integrin is also present in healthy aorta but to a lesser extent.

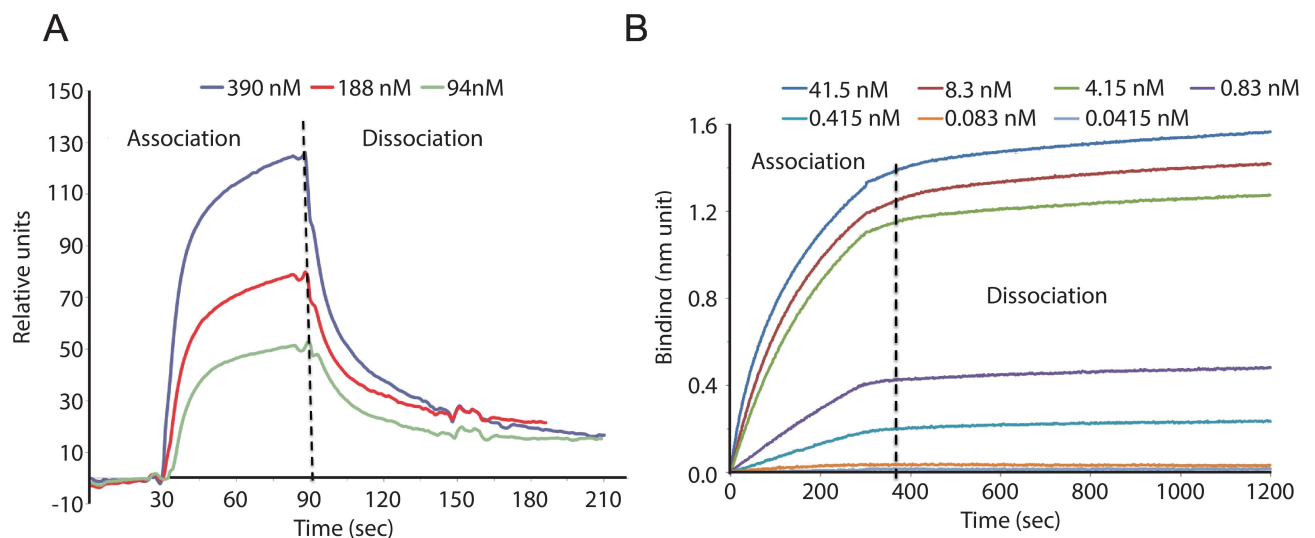
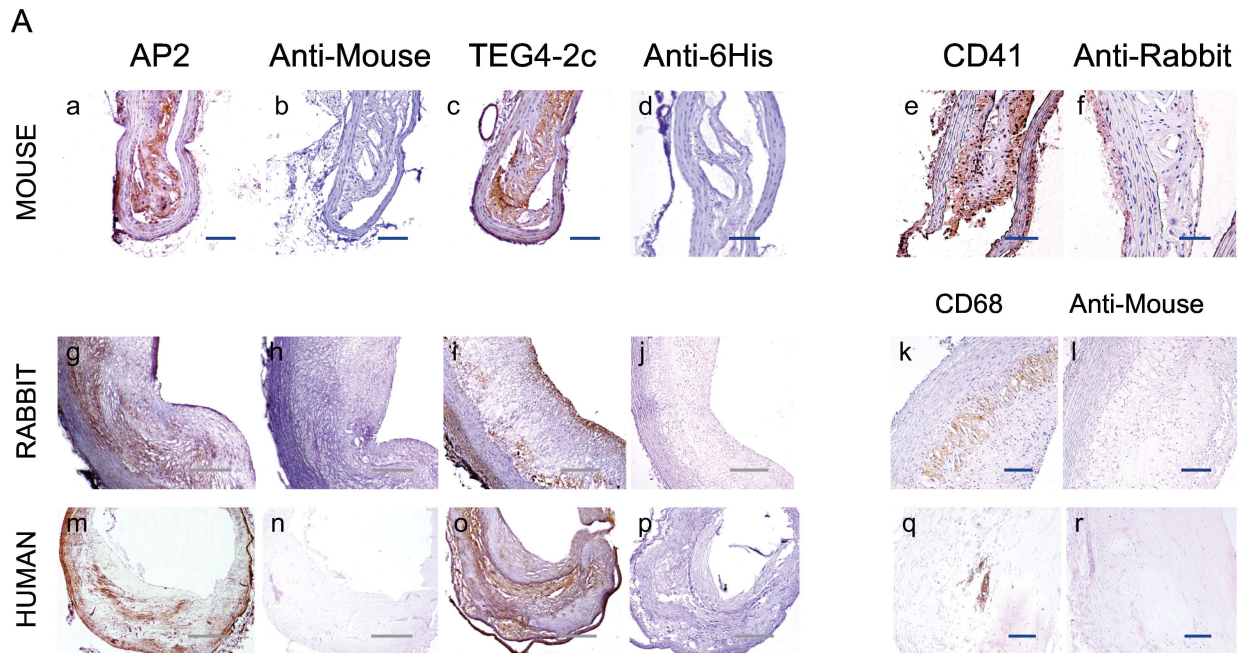


Fig 5. Binding of scFv TEG4-2c to α IIb β 3 by SPR and to whole platelets by BLI. A: SPR sensorgrams. The ligand α IIb β 3 was immobilized on CM5 chip by amine coupling with a density of 8000 RU. Serial dilutions of TEG4-2c in HBS-EP running buffer were injected over the ligand corresponding to 94, 188 and 390 nM. Sensorgrams show a binding concentration-dependent of TEG4-2c scFv. **B: BLI analysis.** TEG4-2c scFv (ligand) was loaded on HIS2 biosensor (anti-penta Histidine Ab) at 21 μ g/mL. Whole platelets (analyte) concentrations converted into α IIb β 3 molarities were: 41.5, 8.3, 4.15, 0.83, 0.415, 0.083 and 0.0415 nM. Additionally one sensor pair was used to record the buffer reference signals. TEG4-2c scFv reacts with α IIb β 3 in its natural conformation in a concentration dependent manner.

doi:10.1371/journal.pone.0170305.g005



B

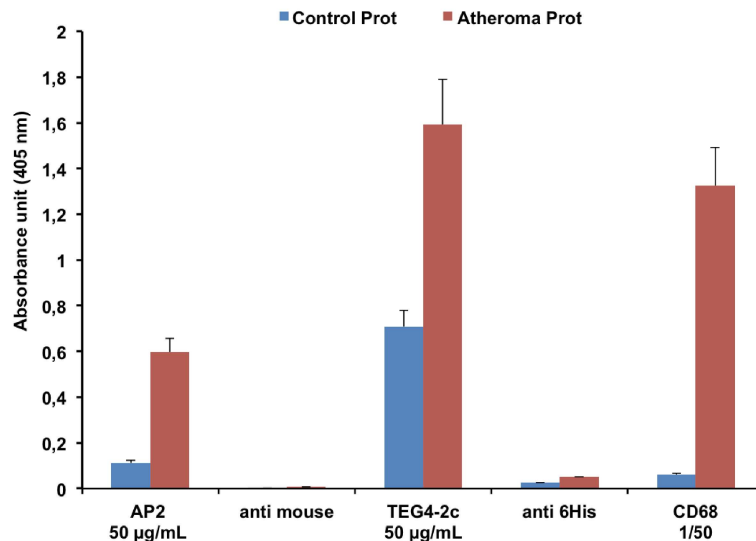


Fig 6. Comparison of the immunoreactivity of TEG4-2c scFv to atherosclerotic tissues of different species by IHC analysis and ELISA assays. **A (a-r):** IHC assays on atherosclerotic tissues: similarly to positive controls e.g; anti-CD41 (anti- α IIb) (e), RAM11 and PGM1 (anti-CD68 antibodies targeting rabbit and human macrophages respectively) (k, q) and AP2 (anti- α IIb β 3 antibody) (a; g; m), TEG4-2c scFv specifically recognizes the injured areas of the aorta sections from different species (c; i; o). Binding of antibodies was visualized via HRP-anti-6His (scFv); HRP anti-rabbit IgG (CD41) and HRP anti-mouse IgG (RAM11, AP2). Negative controls were secondary antibody only (b; d; f; h; j; l; n; p; r). Nuclei were counterstained with hematoxylin **B:** ELISA tests on atheromatous and healthy aorta proteins: TEG4-2c shows a better recognition of atheromatous proteins. RAM11 and AP2 were used as positive controls. Negative controls were secondary antibody only. Binding of antibodies was visualized via HRP-6His or HRP-anti-mouse IgG. OD value represents absorbance at 450 nm. Values represent mean (n = 3) \pm SD (error bars materialized the SD)

doi:10.1371/journal.pone.0170305.g006

Discussion & Conclusion

In the present study, TEG4-2c scFv was expressed at high-level in *Pichia pastoris* using a fed-batch fermentation system monitored by pO₂ level. We produced the TEG4 scFv with cysteine tags at the end of the C-terminal sequence for site-specific conjugation to contrast agents, precluding the loss of reactivity potentially occurring when the grafting process affects antigen-recognition sites. TEG4 scFv had been previously expressed in *E. coli*. [27] Unfortunately, despite optimization tests leading to high yields of cytoplasmic production, proteins also frequently accumulated into inclusion bodies (data not shown). In bacterial systems, many scFv can be produced into the periplasmic space but they are obtained with a very low yield. Higher levels of production can be achieved in inclusion bodies, with the limitation of the presence of insoluble scFv aggregates and the need for subsequent *in vitro* folding that make the use of this bacterial system not attractive for the large scale production of scFv. In addition, many authors have described that the final yield of scFv was only a small percentage of produced proteins with a low specificity for targets. [33,34]

So we chose to use the *Pichia* yeast as an alternative expression system. Indeed, *Pichia pastoris* is an attractive system for low cost-effective large-scale production of heterologous proteins. This type of production, characterized by the secretion of the protein of interest into the culture medium, can be easily scaled up and reach a GRAS (Generally Recognized As Safe) status. The concentration of highly pure TEG4-2c scFv obtained after one chromatography step (IMAC) was up to 600 µg/mL. A production of 30 mg scFv per liter of culture was achieved. This value was lower than that obtained for an scFv anti-carcinoembryonic antigen, [33] but in accordance with the yields obtained for the majority of Fab and antibody fragments expressed in *P. pastoris* [35–37] and sufficient for grafting purposes.

This new scFv format with cysteines included for grafting purposes was tested by ELISA, cytometry and IHC and all the experiments concurred to a specific recognition of platelets, from human or animal model origin and of atheroma issuing from coronary biopsies or animal lesional tissues. Bio-layer interferometry was used for evaluating the affinity of TEG4-2c scFv against platelets because this approach is more relevant than SPR analysis on purified antigen to mimic the *in vivo* behavior. No real K_d value could be extracted from the curves because of an absence of dissociation. This could be explained by the rebinding of the same platelet on the immobilized ligand because of the presence of 50 000 αIIbβ₃ per platelet. Indeed, when the platelet is captured on the surface, one αIIbβ₃ can dissociate while another re-binds on the scFv. As a consequence the off-rate is slower than it would be by using isolated αIIbβ₃. Nevertheless, the bio-layer interferometry underlines a good affinity of TEG4-2c scFv for platelets with recognition in the nanomolar range. Moreover, the flow cytometry results with PRP are in favor of a preferential recognition of activated platelets. This approach, compared to ELISA assays or cytometry performed on washed platelets is the only one that allows recognition in physiological conditions.

The rationale for targeting activated platelets is that they are highly trapped within atherosclerotic lesions not only in thrombi and intraplaque hemorrhage but also in the atheroma burden, around necrotic areas and neovessels, mainly because of the presence of leaky vessels, blood extravasation and haemorrhage. [18] Preferential recognition of activated platelets is to be considered to avoid *in vivo* elimination of the probe by circulating resting platelets.

An important point to be underlined is that scFv TEG4-2c recognizes human platelets as well as murine or rabbit platelets. The use of a human antibody able to target epitopes on the integrin shared by animal models and humans is of particular interest for pre-clinical studies in animal models of the disease and for direct transfer into the clinic. The use of an antibody of human origin is appealing, considering the need for repeated injections in humans to follow

the progression of the pathology. Indeed, the human scFv should limit the induction of anti-antibodies in patients as compared with murine origin; it also avoids any additional humanization step that could impair the reactivity. This antibody equipped with cysteine tags is now ready to be used to functionalize contrast agents for MRI to serve as imaging agents for atherosclerosis.

In conclusion, even if the production of antibody fragments by *P. pastoris* is not always a success story,[38] our findings demonstrated that *P. pastoris* provided TEG4-2c scFv at a high concentration without aggregates and claiming a good purity. Its functionality against activated platelets and atheroma tissues was proven, paving the way for the success of grafting tests. Our results definitely point out the *P. pastoris* expression system as an adapted tool to produce scFv for use in diagnostic or therapeutic applications as well.

Acknowledgments

Research support: This study was funded by (1) the French National Research Agency Program ANR-07-PCVI-0023 IMATHABIO and ANR-13-BSV5-0018 SVSE5 program named ATHERANOS, (2) the SFR TecSan (Technology for Health) of University of Bordeaux and (3) a public grant from the French National Research Agency within the context of the Investments for the Future Program, referenced ANR-10-LABX-0057 and named TRAIL. We would like to thank Arnaud Vonarburg (from Pall Life Sciences) for support on the Octet Red experiments.

Author Contributions

Conceptualization: MJJV GCS AN.

Formal analysis: AVC ML A. Hocquellet JLT XS GCS MJJV AN.

Funding acquisition: GCS XS AN.

Investigation: AVC ML A. Hocquellet A. Hemadou SNP JLT XS GCS MJJV AN.

Methodology: AVC ML A. Hocquellet A. Hemadou SNP JLT XS GCS MJJV AN.

Project administration: GCS XS.

Resources: AVC ML A. Hocquellet A. Hemadou SNP JLT XS GCS MJJV AN.

Supervision: MJJV AN.

Validation: MJJV GCS AN.

Writing – original draft: GCS MJJV AN.

Writing – review & editing: AVC ML A. Hocquellet SNP JLT XS GCS MJJV AN.

References

1. Libby P, Ridker PM, Hansson GK (2011) Progress and challenges in translating the biology of atherosclerosis. *Nature* 473: 317–325. doi: [10.1038/nature10146](https://doi.org/10.1038/nature10146) PMID: [21593864](https://pubmed.ncbi.nlm.nih.gov/21593864/)
2. Falk E (2006) Pathogenesis of atherosclerosis. *J Am Coll Cardiol* 47: C7–12. doi: [10.1016/j.jacc.2005.09.068](https://doi.org/10.1016/j.jacc.2005.09.068) PMID: [16631513](https://pubmed.ncbi.nlm.nih.gov/16631513/)
3. Joshi FR, Lindsay AC, Obaid DR, Falk E, Rudd JH (2012) Non-invasive imaging of atherosclerosis. *Eur Heart J Cardiovasc Imaging* 13: 205–218. doi: [10.1093/ehjci/er319](https://doi.org/10.1093/ehjci/er319) PMID: [22277118](https://pubmed.ncbi.nlm.nih.gov/22277118/)
4. Wildgruber M, Swirski FK, Zernecke A (2013) Molecular imaging of inflammation in atherosclerosis. *Theranostics* 3: 865–884. doi: [10.7150/thno.5771](https://doi.org/10.7150/thno.5771) PMID: [24312156](https://pubmed.ncbi.nlm.nih.gov/24312156/)

5. Weber C, Zernecke A, Libby P (2008) The multifaceted contributions of leukocyte subsets to atherosclerosis: lessons from mouse models. *Nat Rev Immunol* 8: 802–815. doi: [10.1038/nri2415](https://doi.org/10.1038/nri2415) PMID: [18825131](https://pubmed.ncbi.nlm.nih.gov/18825131/)
6. Fenyo IM, Gafencu AV (2013) The involvement of the monocytes/macrophages in chronic inflammation associated with atherosclerosis. *Immunobiology* 218: 1376–1384. doi: [10.1016/j.imbio.2013.06.005](https://doi.org/10.1016/j.imbio.2013.06.005) PMID: [23886694](https://pubmed.ncbi.nlm.nih.gov/23886694/)
7. Burioni R, Canducci F, Saita D, Perotti M, Mancini N, De Marco D, et al. (2009) Antigen-driven evolution of B lymphocytes in coronary atherosclerotic plaques. *J Immunol* 183: 2537–2544. doi: [10.4049/jimmunol.0901076](https://doi.org/10.4049/jimmunol.0901076) PMID: [19635916](https://pubmed.ncbi.nlm.nih.gov/19635916/)
8. Kyaw T, Cui P, Tay C, Kanellakis P, Hosseini H, Liu E, et al. (2013) BAFF receptor mAb treatment ameliorates development and progression of atherosclerosis in hyperlipidemic ApoE(-/-) mice. *PLoS One* 8: e60430. doi: [10.1371/journal.pone.0060430](https://doi.org/10.1371/journal.pone.0060430) PMID: [23560095](https://pubmed.ncbi.nlm.nih.gov/23560095/)
9. Kyaw T, Tay C, Hosseini H, Kanellakis P, Gadowski T, MacKay F, et al. (2012) Depletion of B2 but not B1a B cells in BAFF receptor-deficient ApoE mice attenuates atherosclerosis by potentially ameliorating arterial inflammation. *PLoS One* 7: e29371. doi: [10.1371/journal.pone.0029371](https://doi.org/10.1371/journal.pone.0029371) PMID: [22238605](https://pubmed.ncbi.nlm.nih.gov/22238605/)
10. Kyaw T, Tay C, Khan A, Dumouchel V, Cao A, To K, et al. (2010) Conventional B2 B cell depletion ameliorates whereas its adoptive transfer aggravates atherosclerosis. *J Immunol* 185: 4410–4419. doi: [10.4049/jimmunol.1000033](https://doi.org/10.4049/jimmunol.1000033) PMID: [20817865](https://pubmed.ncbi.nlm.nih.gov/20817865/)
11. Herter JM, Rossaint J, Zarbock A (2014) Platelets in inflammation and immunity. *J Thromb Haemost* 12: 1764–1775. doi: [10.1111/jth.12730](https://doi.org/10.1111/jth.12730) PMID: [25224706](https://pubmed.ncbi.nlm.nih.gov/25224706/)
12. Morrell CN, Aggrey AA, Chapman LM, Modjeski KL (2014) Emerging roles for platelets as immune and inflammatory cells. *Blood* 123: 2759–2767. doi: [10.1182/blood-2013-11-462432](https://doi.org/10.1182/blood-2013-11-462432) PMID: [24585776](https://pubmed.ncbi.nlm.nih.gov/24585776/)
13. Lievens D, Zernecke A, Seijkens T, Soehnlein O, Beckers L, Munnix IC, et al. (2010) Platelet CD40L mediates thrombotic and inflammatory processes in atherosclerosis. *Blood* 116: 4317–4327. doi: [10.1182/blood-2010-01-261206](https://doi.org/10.1182/blood-2010-01-261206) PMID: [20705757](https://pubmed.ncbi.nlm.nih.gov/20705757/)
14. Ferdous F, Scott TR (2015) A comparative examination of thrombocyte/platelet immunity. *Immunol Lett* 163: 32–39. doi: [10.1016/j.imlet.2014.11.010](https://doi.org/10.1016/j.imlet.2014.11.010) PMID: [25448707](https://pubmed.ncbi.nlm.nih.gov/25448707/)
15. Aloui C, Prigent A, Sut C, Tariket S, Hamzeh-Cognasse H, Pozzetto B, et al. (2014) The signaling role of CD40 ligand in platelet biology and in platelet component transfusion. *Int J Mol Sci* 15: 22342–22364. doi: [10.3390/ijms151222342](https://doi.org/10.3390/ijms151222342) PMID: [25479079](https://pubmed.ncbi.nlm.nih.gov/25479079/)
16. Daub K, Seizer P, Stellos K, Kramer BF, Bigalke B, Schaller M, et al. (2010) Oxidized LDL-activated platelets induce vascular inflammation. *Semin Thromb Hemost* 36: 146–156. doi: [10.1055/s-0030-1251498](https://doi.org/10.1055/s-0030-1251498) PMID: [20414829](https://pubmed.ncbi.nlm.nih.gov/20414829/)
17. Badrnya S, Schrottmaier WC, Kral JB, Yaiw KC, Volf I, Schabbauer G, et al. (2014) Platelets mediate oxidized low-density lipoprotein-induced monocyte extravasation and foam cell formation. *Arterioscler Thromb Vasc Biol* 34: 571–580. doi: [10.1161/ATVBAHA.113.302919](https://doi.org/10.1161/ATVBAHA.113.302919) PMID: [24371083](https://pubmed.ncbi.nlm.nih.gov/24371083/)
18. Jacobin-Valat MJ, Laroche-Traineau J, Lariviere M, Mornet S, Sanchez S, Biran M, et al. (2015) Nanoparticles functionalised with an anti-platelet human antibody for in vivo detection of atherosclerotic plaque by magnetic resonance imaging. *Nanomedicine* 11: 927–937. doi: [10.1016/j.nano.2014.12.006](https://doi.org/10.1016/j.nano.2014.12.006) PMID: [25684334](https://pubmed.ncbi.nlm.nih.gov/25684334/)
19. Reichert JM (2016) Antibodies to watch in 2016. *MAbs* 8: 197–204. doi: [10.1080/19420862.2015.1125583](https://doi.org/10.1080/19420862.2015.1125583) PMID: [26651519](https://pubmed.ncbi.nlm.nih.gov/26651519/)
20. Frenzel A, Hust M, Schirrmann T (2013) Expression of recombinant antibodies. *Front Immunol* 4: 217. doi: [10.3389/fimmu.2013.00217](https://doi.org/10.3389/fimmu.2013.00217) PMID: [23908655](https://pubmed.ncbi.nlm.nih.gov/23908655/)
21. Frenzel A, Schirrmann T, Hust M (2016) Phage display-derived human antibodies in clinical development and therapy. *MAbs* 8: 1177–1194. doi: [10.1080/19420862.2016.1212149](https://doi.org/10.1080/19420862.2016.1212149) PMID: [27416017](https://pubmed.ncbi.nlm.nih.gov/27416017/)
22. Groves MA, Amanuel L, Campbell JL, Rees DG, Sridharan S, Finch DK, et al. (2014) Antibody VH and VL recombination using phage and ribosome display technologies reveals distinct structural routes to affinity improvements with VH-VL interface residues providing important structural diversity. *MAbs* 6: 236–245. doi: [10.4161/mabs.27261](https://doi.org/10.4161/mabs.27261) PMID: [24256948](https://pubmed.ncbi.nlm.nih.gov/24256948/)
23. Ahmad ZA, Yeap SK, Ali AM, Ho WY, Alitheen NB, Hamid M (2012) scFv antibody: principles and clinical application. *Clin Dev Immunol* 2012: 980250. doi: [10.1155/2012/980250](https://doi.org/10.1155/2012/980250) PMID: [22474489](https://pubmed.ncbi.nlm.nih.gov/22474489/)
24. Jacobin MJ, Laroche-Traineau J, Little M, Keller A, Peter K, Welschof M, et al. (2002) Human IgG monoclonal anti-alpha(IIB)beta(3)-binding fragments derived from immunized donors using phage display. *J Immunol* 168: 2035–2045. PMID: [11823541](https://pubmed.ncbi.nlm.nih.gov/11823541/)
25. Hagemeyer CE, von Zur Muhlen C, von Elverfeldt D, Peter K (2009) Single-chain antibodies as diagnostic tools and therapeutic agents. *Thromb Haemost* 101: 1012–1019. PMID: [19492141](https://pubmed.ncbi.nlm.nih.gov/19492141/)
26. Hohmann JD, Wang X, Krajewski S, Selan C, Haller CA, Straub A, et al. (2013) Delayed targeting of CD39 to activated platelet GPIIb/IIIa via a single-chain antibody: breaking the link between

- antithrombotic potency and bleeding? *Blood* 121: 3067–3075. doi: [10.1182/blood-2012-08-449694](https://doi.org/10.1182/blood-2012-08-449694) PMID: [23380744](https://pubmed.ncbi.nlm.nih.gov/23380744/)
27. Robert R, Noubhani AM, Jacobin MJ, Santarelli X, Clofent-Sanchez G (2005) Improvement in production and purification bioprocesses of bacterially expressed anti-alphaIIb beta3 human single-chain Fv antibodies. *J Chromatogr B Analyt Technol Biomed Life Sci* 818: 43–51. doi: [10.1016/j.jchromb.2004.10.038](https://doi.org/10.1016/j.jchromb.2004.10.038) PMID: [15722043](https://pubmed.ncbi.nlm.nih.gov/15722043/)
 28. Narasimhan Janakiraman V, Noubhani A, Venkataraman K, Vijayalakshmi M, Santarelli X (2016) High yield of recombinant human Apolipoprotein A-I expressed in *Pichia pastoris* by using mixed-mode chromatography. *Biotechnol J* 11: 117–126. doi: [10.1002/biot.201500245](https://doi.org/10.1002/biot.201500245) PMID: [26626736](https://pubmed.ncbi.nlm.nih.gov/26626736/)
 29. Robert R, Clofent-Sanchez G, Hocquellet A, Jacobin-Valat MJ, Daret D, Noubhani AM, et al. (2006) Large-scale production, bacterial localization assessment and immobilized metal affinity chromatography purification of a human single-chain Fv antibody against alphaIIb-beta3 integrin. *Int J Biol Macromol* 39: 51–59. doi: [10.1016/j.ijbiomac.2006.01.014](https://doi.org/10.1016/j.ijbiomac.2006.01.014) PMID: [16620955](https://pubmed.ncbi.nlm.nih.gov/16620955/)
 30. Dabadie M, Valli N, Jacobin MJ, Laroche-Traineau J, Barat JL, Ducassou D, et al. (2001) Characterisation, cloning and sequencing of a conformation-dependent monoclonal antibody to the alphaIIb beta3 integrin: interest for use in thrombus detection. *Platelets* 12: 395–405. doi: [10.1080/09537100120071031](https://doi.org/10.1080/09537100120071031) PMID: [11674856](https://pubmed.ncbi.nlm.nih.gov/11674856/)
 31. Shattil SJ, Hoxie JA, Cunningham M, Brass LF (1985) Changes in the platelet membrane glycoprotein IIb/IIIa complex during platelet activation. *J Biol Chem* 260: 11107–11114. PMID: [2411729](https://pubmed.ncbi.nlm.nih.gov/2411729/)
 32. Pidard D, Montgomery RR, Bennett JS, Kunicki TJ (1983) Interaction of AP-2, a monoclonal antibody specific for the human platelet glycoprotein IIb-IIIa complex, with intact platelets. *J Biol Chem* 258: 12582–12586. PMID: [6226659](https://pubmed.ncbi.nlm.nih.gov/6226659/)
 33. Freyre FM, Vazquez JE, Ayala M, Canaan-Haden L, Bell H, Rodriguez I, et al. (2000) Very high expression of an anti-carcinoembryonic antigen single chain Fv antibody fragment in the yeast *Pichia pastoris*. *J Biotechnol* 76: 157–163. PMID: [10656330](https://pubmed.ncbi.nlm.nih.gov/10656330/)
 34. Joosten V, Lokman C, Van Den Hondel CA, Punt PJ (2003) The production of antibody fragments and antibody fusion proteins by yeasts and filamentous fungi. *Microb Cell Fact* 2: 1. doi: [10.1186/1475-2859-2-1](https://doi.org/10.1186/1475-2859-2-1) PMID: [12605725](https://pubmed.ncbi.nlm.nih.gov/12605725/)
 35. Eldin P, Pauza ME, Hieda Y, Lin G, Murtaugh MP, Pentel PR, et al. (1997) High-level secretion of two antibody single chain Fv fragments by *Pichia pastoris*. *J Immunol Methods* 201: 67–75. PMID: [9032410](https://pubmed.ncbi.nlm.nih.gov/9032410/)
 36. Takahashi K, Yuuki T, Takai T, Ra C, Okumura K, Yokota T, et al. (2000) Production of humanized Fab fragment against human high affinity IgE receptor in *Pichia pastoris*. *Biosci Biotechnol Biochem* 64: 2138–2144. doi: [10.1271/bbb.64.2138](https://doi.org/10.1271/bbb.64.2138) PMID: [11129586](https://pubmed.ncbi.nlm.nih.gov/11129586/)
 37. Lange S, Schmitt J, Schmid RD (2001) High-yield expression of the recombinant, atrazine-specific Fab fragment K411B by the methylotrophic yeast *Pichia pastoris*. *J Immunol Methods* 255: 103–114. PMID: [11470291](https://pubmed.ncbi.nlm.nih.gov/11470291/)
 38. Cupit PM, Whyte JA, Porter AJ, Browne MJ, Holmes SD, Harris WJ, et al. (1999) Cloning and expression of single chain antibody fragments in *Escherichia coli* and *Pichia pastoris*. *Lett Appl Microbiol* 29: 273–277. PMID: [10664964](https://pubmed.ncbi.nlm.nih.gov/10664964/)

A la recherche d'anticorps humains et détermination de nouveaux biomarqueurs pour le ciblage de la plaque d'athérome

Résumé : Cette thèse porte sur la sélection d'anticorps humains et la découverte de nouveaux biomarqueurs en vue d'un diagnostic et d'une thérapie de l'athérosclérose. Le premier volet concerne, dans un premier temps, la sélection d'anticorps humains dans un contexte physiopathologique chez un modèle animal de l'athérosclérose. La sélection *in vivo* a permis de sélectionner des anticorps sur large panel de cibles dans leur conformation native et sous l'influence de leur microenvironnement. Suite à cela, afin d'accéder à la spécificité des clones, une méthode innovante de criblage a été développée, la cytométrie de flux à moyen débit. Cette technique, contrairement à l'ELISA couramment utilisé pour la caractérisation d'anticorps ou de peptides, ne nécessite pas de connaître l'antigène et permet également de travailler à des concentrations de protéines moindres. Ainsi, deux cents anticorps ont pu être sélectionnés. Les tests d'affinité et de caractérisation de ces anticorps ont, par la suite, mené à l'identification d'une nouvelle cible de la plaque d'athérome, la galectine 3, protéine impliquée dans l'activation des macrophages ainsi que la rétention des LDL oxydées par ces dernières. Le deuxième volet concerne une seconde approche de criblage *in silico* des anticorps issus de la sélection *in vivo*, notamment grâce à l'utilisation de techniques de séquençage haut débit (Pacific Biosciences). Les résultats de cette étude pilote ont permis de mettre en exergue plusieurs points : (1) la possibilité d'obtenir des séquences >1000bp correspondant à des scFv après analyse par IMGT/HighV-QUEST (database des immunoglobulines internationales) et (2) la possibilité d'identifier des séquences de clones préférentiellement enrichies après une sélection *in vivo*. En conclusion et perspectives, ces deux méthodes de criblage complémentaires nous donnent accès à de nouvelles voies d'obtention d'anticorps humains pour (1) le couplage à des nanoparticules pour le diagnostic par IRM de la plaque d'athérome ainsi que (2) apporter une thérapie *in situ* au site inflammatoire.

Discovery of human antibodies and identification of new biomarkers for atheroma targeting

Summary: This thesis is based on the selection of human antibodies and identification of new biomarkers to develop a diagnosis and a therapy of atherosclerosis. The part one treats, first, the selection of human antibodies in a pathophysiological context in an animal model of atherosclerosis. To access not only to a large panel of targets but also proteins in their native conformation and influenced by their microenvironment, an *in vivo* selection was performed. In a second time, an innovative high throughput flow cytometry screening was implemented. Contrary to the use of ELISA screening which needs to have an identified target, the flow cytometry doesn't have this limitation and uses a lower quantity of proteins than ELISA assays. By this screening, two thousand clones were selected. Characterization and specificity tests of these antibodies have yielded to the discovery of galectin 3 has a new biomarker. This protein is implicated in macrophage activation, binding of oxLDL by macrophages to form foam cells. The second part is about a second screening approach of the antibodies issuing from the *in vivo* selection by using Next Generation Sequencing methods (Pacific Biosciences). The results of this primary study allowed for highlighting two points: (1) the feasibility to obtain long reads >1000bp identified as scFv after analysis by IMGT/HighV-QUEST (international database of immunoglobulins) and (2) the feasibility to identify hits of clones preferentially enriched during the *in vivo* selection. In conclusion and perspectives, these two complementary methods give the opportunity to obtain human antibodies for (1) the coupling to nanoparticles for diagnosis by MRI of atheroma plaque and (2) serve as ligands for *in situ* delivery of drugs.