IDENTIFY - Identify, Characterize and Measure Biomolecules in a Variety of Sample Sources

Biomarker Discovery and Validation by HPLC Chip-LC/MS

Keith Waddell LC/MS Applications Manager October 22, 2008





#### **Overview**

HPLC-Chip/MS for reproducible, robust nanoflow LC/MS

- Biomarker discovery using accurate mass with a mass profiling approach
- Advanced statistical analysis using GeneSpring MS
- Biomarker validation using HPLC-Chip/QQQ for quantitative measurements



## Why Use Microfluidics For Nanospray LC/MS?

Integrate functional components onto a reusable, biocompatible chip

- enrichment and analytical nanocolumns,
- nanospray emitter
- fittings and connection capillaries
- directly on a reusable biocompatible polymer chip.





## HPLC-Chip/MS Interface: Fluid Connections to the HPLC-Chip





## **Better Chromatography Reduces Interferences in Complex Matrices**





## **Retention Time Reproducibility**



Extracted ion chromatograms for 17 peaks from a BSA tryptic digest (50 fmol on-column)

	RT	SD	%RSD
EIC 487.8	3.618	0.014	0.40
EIC 752	3.788	0.011	0.29
EIC 740.6	5.018	0.010	0.20
EIC 874.4	3.968	0.012	0.31
EIC 653.6	4.289	0.012	0.28
EIC 511.7	3.681	0.012	0.31
EIC 722.7	3.547	0.012	0.35
EIC 778	4.143	0.010	0.23
EIC 526.3	4.399	0.015	0.34
EIC 547.5	4.472	0.011	0.25
EIC 746.7	5.196	0.011	0.20
EIC 519.1	4.142	0.011	0.26
EIC 508.2	4.972	0.011	0.23
EIC 582.4	4.679	0.011	0.23
EIC 461.9	3.905	0.012	0.30
EIC 474	4.759	0.011	0.22
EIC 628	4.584	0.010	0.22

RT reproducibility evaluated using 69 repeat injections



### Mass Profiling- Find The Differences Between Samples



Molecular Feature: a discrete molecular entity defined by the combination of

- retention time, mass and response in an LC/MS analysis
- retention time, mass spectrum and response in a GC/MS analysis



• Find real differences in sample sets using statistical analysis

• Reproducible measurements minimize the number of samples!



# MS Profiling And MS/MS Identification With The Agilent 6520 Q-TOF And HPLC-Chip





### **Mass Profiling Software**





### **GeneSpring MS for Biomarker Discovery**



GeneSpring MS



#### **Candidate Identification From Targeted MS/MS**

Back -	Favorites Lools	Help	Favorites	<i>6</i>	- 🛃 👿	- 🛄 🕯	8	
Address 🙆 http://sp	ectrummilletd/millht	ml/summaryframe.htm			Jac.			Con Links ** 📆 🗸
Agilent Spectrur	n Mill - Protein	/Peptide Summa	у					
Spectrum Mill	Summary S	ettings Autovali	dation Bui	ild TIC MS	5/MS Searc	h Spectr	rum Sur	mmary Tool Belt Help
30min # spectra mean intensity	60min # spectra mean intensity	Archakov\\SUX- frs # spectra mean intensity	Database Accession #	%AA Coverage	Distinct Peptides (#)	Summed MS/MS Search Score	Group #	Protein Name
4 1.11e+006	12 1.08e+007	28 1.24e+006	P20029	<u>39</u>	21	319.95	1.1	2X 78 kDa glucose-regulated protein precursor (GRP 78) (Immunoglobulin heavy chain-binding protein) (BiP)
9 2.23e+006	14 1.02e+007	22 2.12e+006	<u>P09103</u>	<u>38</u>	18	261.60	2.1	Protein disulfide-isomerase precursor (EC 5.3.4.1) (PDI) (Prolyl 4-hydroxylase beta subunit) (Cellular thyroid hormone-bindin
9 3.26e+006	27 1.10e+007	33 1.72e+006	<u>Q64458</u>	<u>39</u>	14	215.05	3.1	3X Cytochrome P450 2C29 (EC 1.14.14.1) (CYPIIC29) (P-450 MUT-2) (Aldehyde oxygenase)
6 1.86e+006	16 8.35e+006	12 2.50e+006	<u>Q8∨CT4</u>	<u>33</u>	13	201.81	<u>4.1</u>	4 X Carboxylesterase 3 precursor (EC 3.1.1.1) (Triacylglycerol hydrolase) (TGH)
2 3.76e+005	11 1.48e+007	19 2.40e+006	<u>Q64459</u>	<u>31</u>	13	196.62	<u>5.1</u>	Cytochrome P450 3A11 (EC 1.14.14.1) (CYPIIIA11) (P-450IIIAM1) (P-450UT)
10 3.66e+006	11 1.38e+007	20 1.76e+006	<u>Q63880</u>	<u>34</u>	13	195.98	<u>6.1</u>	Liver carboxylesterase 31 precursor (EC 3.1.1.1) (ES-Male) (Esterase-31) 💌
3 6.26e+005	4 5.82e+006	14 1.23e+006	<u>P37040</u>	22	10	154.70	<u>7.1</u>	NADPH-cytochrome P450 reductase (EC 1.6.2.4) (CPR) (P450R)
1 7.15e+005	5 1.13e+007	12 8.57e+005	<u>P27773</u>	20	10	142.58	<u>8.1</u>	Protein disulfide-isomerase A3 precursor (EC 5.3.4.1) (Disulfide isomerase ER-60) (ERp60) (58 kDa microsomal protein) (p5
5 2.96e+006	10 8.80e+006	9 1.81e+006	<u>P24456</u>	27	9	140.41	<u>9.1</u>	3X Cytochrome P450 2D10 (EC 1.14.14.1) (CYPIID10) (P450-16-alpha) (P450CB) (Testosterone 16-alpha hydroxylase)
7 4.17e+006	19 1.36e+007	21 1.73e+006	<u>Q63886</u>	<u>25</u>	9	136.88	<u>10.1</u>	ZX UDP-glucuronosyltransferase 1-1 precursor, microsomal (EC 2.4.1.17) (UDPGT) (UGT1*1) (UGT1-01) (UGT1.1) (UGT1A)
0 0.00e+000	4 7.76e+006	14 1.21e+006	<u>Q9D379</u>	23	10		March Million Holman Fac	2.3) (Microsomal epoxide hydrolase) (Epoxide hydratase) 💌
3 4.57e+006	5 1.14e+007	10 1.68e+006	<u>Q62397</u>	22	9	an all ann santan ann an Ona - O 2 2 2 0 € Santan Santanan Aglant Santanangan Sg	ectrum Mill MI Profe	III.1) (CYPIB10) (Testosterone 16-alpha hydroxylase) (P450-16-alpha) (Clone PF3/46) 🗸
3 2.72e+005	6 1.05e+007	8 2.17e+006	<u>P08113</u>	<u>13</u>	9		A Delevator - Dele	Ismic reticulum protein 99) (94 kDa glucose-regulated protein) (GRP94) (ERP99) (Polymorp
5 7.29e+005	2 6.83e+006	3 1.06e+006	<u>008601</u>	<u>11</u>	8	Party Republic Will Bases Million Reach Party Descent Party Descent Party Descent Reach Descent Reac	Floored) De Interneting Tester Intel Payette Connessey Se estuary Commany Se intel Commany Se intel Commany Se	en lasca de la contraction de
9 8.90e+005	7 7.21e+006	14 7.13e+005	<u>Q8JZR0</u>	<u>16</u>	7	Tipe and Troks FADs Adds Totals Table Real State Real Mathematics Real Dipyther Reason Page	nitet Co nititations Hi 47C Di Ala bente Sa	as representation and of the provide state of the state
3 1.12e+006	2 6.06e+006	6 1.35e+006	<u>Q05421</u>	<u>14</u>	7	Records May Code Allow Allow Allow Allow Code State Code Allow All	ogan Pr Steven St Prokat Pr Comp O Prologe D	1.1) (CYPIIE1) (P450-J) (P450-ALC)
· ··· <sup>0</sup> ····	5	4	P08003	15	7	ECompl 202 203 Mind Telescope to Ecompl 202 All All All All All All All All All Al	ndi Jir Heli d'Laffrens Ni	ecursor (EC 5.3.4.1) (Protein ERo-72) (ERo72)
							otr	



### Protein Biomarker Model Study on a Complex Mixture

Five mixtures prepared which contained:

- Tryptic digest of an *E. coli* lysate (complex background)
- Tryptic digests of bovine serotransferrin and BSA at specified levels
   Each mixture was prepared 10 times
- One injection of each sample (1  $\mu$ L injection volume)
- Long (100 min injection-to-injection) method used



## Complex Model Study: Spiked E. coli Lysate

Sample	<i>E. coli</i> lysate (ng total protein)	BSA (fmol)	Serotransferrin (fmol)
1	400	25	200
2	400	50	100
3	400	100	50
4	400	200	25
5	400	400	5



# Complex Model Study: Spiked *E. coli* Lysate Total Ion Chromatogram



- Digest of *E. coli* lysate creates a complex background of thousands of peptides
- Profiling software must be able to find the bovine peptides that were spiked into the *E. coli* mixture



#### **Complex Model Study: Molecular Feature Extraction**

#### Raw Data



#### **Extracted Features**





### **Complex Model Study: Reproducibility With Technical Replicates**



Color-coding by molecular feature

Total of 762 features shown

Average SD is 0.0519 min for RT 1.6 mDa for mass



## **Complex Model Study: Clustering of Features That Are Significantly Different (ANOVA)**

#### **Cluster by sample**

**Cluster by level** 





#### **Complex Model Study: Finding The Differential Features in GeneSpring MS**





#### **Complex Model Study: K-Means Clustering of Differential Features And PMF Search**

	Sample ID (comr Database searche Molecular weight Full pI range: <b>358</b> Combined molect PMF search select	nent): Enter_Commen ed: SwissProt.mamma search (1000 - 10000 824 entries. ular weight and pI sear cts 25 entries.	nt als O Da) selects ches select 31	Par 31701 entries. 1701 entries.	ameters used i	in Search
Normalized Internaty	Dynamic Rank Probability P Score	Static # (%) robability Masses Score <u>Matched</u>	Mass Error Mean (Std H Dev) C (ppm)	Protein Protein Soverage MW (Da)	/pI Species	Accession # Protein Name
10 <sup>-</sup>	$     \begin{array}{c}             \underline{1} & 2.28e- \\             \underline{008} & 0 \\             \underline{2} & 0.71 & 0 \\             \underline{3} & 1.8 & 1 \\             \underline{4} & 2.76 & 2         \end{array} $	.16e- 07 32/515 (6%) .71 27/515 (5%) .8 13/515 (2%) .76 14/515 (2%)	3.6 (6.6) -2.5 (10.9) -2.8 (11.4) -2.8 (10.9)	49%         77753.7/6.           40%         75755.5/5.           30%         46512.7/6.           28%         41450.6/10	75 BOVIN 34 MOUSE 54 MOUSE 1.65 HUMAN	Q29443       Serotransferrin precursor (Transferrin) (Siderophilin) (Beta-1-metal-binding globulin)         P14824       Annexin A6 (Annexin VI) (Lipocortin VI) (P68) (P70) (Protein III) (Chromobindin-20) (67 kDa ca         Q91WL8       WW domain-containing oxidoreductase (EC 1.1.1)         075683       Surfeit locus protein 6
A difference of the second sec	masse	S	In the second seco	20 15 12 10 00 00 00 00 00 00 00 00 00 00 00 00		Create Inclusion List for QTOF         Select the parameters to create inclusion list         Select RT Tolerance            • Narrow          Select criteria for lon selection         • All Z states         • Most Abundant         • Prefer 2+ above threshold         Select output file         • OK         • Cancel



### **Complex Model Study: Mass Spectra From a Targeted Peptide**



#### Apex mass spectrum for a targeted species



Fragment ion assignments for targeted peptide (from transferrin)



## **Biomarker Discovery Using Protein Profiling:** Summary

- Targeted biomarker discovery can reveal and identify proteins usually missed when using the shotgun data-dependent approach
- Label-free workflow saves cost during discovery
- Biomarker discovery software is essential for identifying candidates
  - Molecular feature extraction algorithm detects low level peaks in complex mixtures
  - MassHunter Profiling software determines differentially expressed features
  - GeneSpring MS software offers a powerful data analysis/visualization tools when analyzing large sample sets
  - Spectrum Mill Protein Identification software provides confident protein results to move onto the next steps of biomarker research



### **Jump From Discovery Phase to Validation Phase**

- Reduce the time needed for analysis
- Increase throughput
- Improve CV
- Reduce cost





	Research	Clinical				
# of samples	<100	Hundreds - thousands				
# of proteins	50-500	1-20				
Time	Months - years	minutes - hours				
Cost	\$100k-1M	\$10-100				
CV	20-50%	3-5%				



#### Agilent HPLC-Chip/QQQ LCMS Technology Nanospray chip configuration brings new era in high sensitivity quantitation



Sensitivity: down to low amol Dynamic range: up to 10<sup>5</sup>



#### Agilent's New Axial Acceleration Collision Cell

Overcomes memory or cross-talk effects!

- using high speed ion transport

Maximum sensitivity

- using wide mass range hexapole design

Simple to operate

- no complicated wave forms





#### **Triple Quadrupole: SRM**





#### **Excellent Reproducibility of MS Response** SRM of HSA Peptide LVNEVTEFAK from 10 amol to 1 pmol (n=6)



All RSDs are within 15%



## Biomarker Discovery and Validation Method Development





### Limit of Quantitation for Peroxidase Spiked in Human Serum: 10 amol to 10 fmol





#### **External Quantitation Curve of Peroxidase Peptide** DTIVNELR From 10 amol - 10 fmol Spiked in Human Serum





## Absolute Protein Quantification in the Context of Non-clinical Drug Safety Evaluation

InnoMed PredTox Consortium: 15 industrial and 3 academic partners

Goal: Assess value of combining "omics" data with traditional toxicology data for preclinical safety evaluation



Collins B. C. et al. ASMS 2008 MPQ 477



#### **Experimental Design**

Catalase was selected based on previous 2D-DIGE data

Rat liver lysate were prepared from rats treated with troglitazone (hepatotoxicant) or vehicle control

Peptides and MRM transitions were selected using Peptide Selector in Spectrum Mill and <sup>13</sup>C, <sup>15</sup>N labeled peptides were synthesized

1 mg of soluble protein extract was reduced, alkylated, acetone precipitated and trypsin digested

The liver digest were spiked with the isotope-labeled peptides and analyzed by Agilent 6410 QQQ system



## Using Spectrum Mill Peptide Selector for Optimizing MRM Transitions

Chemically reactive residues (Cys = C, Met = M, Trp = W)

Residues with variable PTM

Peptides adjacent to multiple cleavage site

Size of the peptides

Uniqueness of the sequence in the database

Peptide Selector - Agilent Spectrum Mill - Microsoft Internet Explorer provided by Agilent Technologies, Inc.	(	
<u>File Edit Vi</u> ew F <u>a</u> vorites <u>T</u> ools <u>H</u> elp		-
🔾 Back + 🕥 - 🖹 😰 🏠 🔎 Search 🬟 Favorites 🤣 😥 - چ 🔟 + 🛄 鑬 🦓		
Address 🕘 http://smapps/millhtml/mssluice.htm	🔽 🄁 Go	Links <b>*</b>
Agilent Spectrum Mill - Peptide Selector		^
Spectrum Mill MS Edman Multiple Sequence Aligner Databases MS Digest Help		
Selection		
Select 🔲 Hide HTML links (better Excel cut/paste)		
Ninest Parameters Product Ion Parameters		
Digest: Trypsin V Maximum # missed cleavages: 0 V Show Product Ion Masses		
Criteria for Excluding Peptides		
Maximum # basic residues (RHK): 1 Y Peptide exclusion criteria: AA Composition Filtering:		
Minimum peptide MH*: 900.0 🛛 🗹 Has nearby cleavage site within 3 🔽 residues Required AAs: KR		
Contains peptide N-terminal GIn to pyroGlu     Disallowed AAs: CM		
Contains protein N-terminus Acetylated		
Contains to variable modification		
Modifications		
Choose Fixed: Carbamidomethylation (C) Variable:		
Protein(s) to Select From Search Mode		
Database: SwissProt Count Peptide Uniqueness in Database by: None		
Select Peptides from only the Database Entries Species: All		
with the accession numbers below Allowed delimiters (.):.)		
P00433		
	62	~
	Second Second Second	



#### **Peptide Selector – Catalase Results**

	Peptide Selector Results	<b></b>
Number of database entries:211104 Database: SwissProt Exclusion criteria:   Peptide N-terminal GIn to pyroGlu   Required AA's: KR Disallowed AA's: CM Exclude if Nearby Cleavage Sites within 3 residues Digest Max. # Missed Fixed Mods Peptide Used Cleavages carbamidomethylation N terminus	Peptide Uniqueness C terminus Masses are Count by	
Trypsin 0 Carbannooneuryladon Hydrogen (H) Protein Name: Catalase (FC 1.11.1.6)	Free Acid (O H) Sequence	
Species: RAT		
SwissProt Accession #: P04762 MS-Digest Index #: 18675		
pl of Protein: 7.15		
Amino Acid Composition: A42 C8 036 E25 F31 G35 H21 I20 K30 L3	1 M12 N33 P3-926 R30 S24 T25 V33 W6 Y21	
Protein Name Acc # RP- MH+ m/2 #	DB Start End eps AA AA Prev. Sequence Next b <sub>2</sub> C slue of Asp, Glu N.side of Pro	
Catalase (EC 1.11.1.6) P04 62 13.63 984.5109 492.7591	<u>1</u> 243 251 (GIK) <u>NLPVEEAGR</u> (LAQ) <b>228.14</b> 870.47 <b>757.38</b> 660.33 561.26 <b>432.22 303.18</b> 232.14	175.1
Catalase (EO 1.11.1.6) P04762 18.97 1001.5666 501.2869	<u>6</u> 306 314 (PHK) <u>DYPLIPVGK</u> (LVL) <b>279.10</b> 886.54 723.48 626.42 513.34 400.26 303.20 204.13	147.1
Catalase (EC 1.11.1.6) P04762 19.22 1276.6168 638.8120	2 252 262 (AGR) <u>LAQEDPDYGLR</u> (DLF) <b>185.13</b> 1163.53 1092.50 964.44 <b>835.39</b> 720.37 623.31 <b>508.29</b>	345.2
Catalase (EC 1.11.1.6) P04762 22.02 1655.7952 828.4012	<u>2</u> 287 300 (TFK) <u>EAETFPFNPFDLTK</u> (VWP) <b>201.09 1526.75</b> 1455.72 <b>1326.67</b> 1225.63 <b>1078.56</b> 981.50 834.44	720.3
Catalase (EC 1.11.1.6) P04762 26.89 2518.2038 1259.6055	<u>7</u> 135 155 (AVK) <u>FYTEDGIVIO VGNNTPIFFIR</u> (DAM) <b>311.14</b> 2371.14 2208.07 2107.02 1977.56 1862.95 1805.93 1691.69 1	1505.8
Catalase (EC 1.11.1.6) P04762	9	
	NTACHDONI NONNETDEM AUFDREDIDE DINNAVES 00	
1 ADSKDEASDQ MAQWAEQAAP QAPDVLIIGG GAPIGDALAI 81 FOVFFVTHDI TRYSKAKVEF HIGKRTDIAU RESTVAGESG	SADTURDERG FAUKEVTEDG NUDI WONNER TEETEDAMIE 160	
161 PSFTHSOKRN POTHLKDPDM VWDFWSLCPE SLHOVTFLFS	DRGIPDGHRH MNGYGSHTEK LVNANGEAVY CKEHYKTDOG 240	
241 IKNLPVEEAG RLAOEDPDYG LRDLFNAIAS GNYPSWTFYI	OVMTFKEAET FPFNPFDLTK VWPHKDYPLI PVGKLVLNRN 320	
321 PANYFAEVEQ MAFDPSNMPP GIEPSPDKML QGRLFAYPDT	303.20 204.13 147.11 HRHRLGPNYL QIPVNCPYRA RVANYQRDGP MCMHDNQGGA 400	
401 PNYYPNSFSA PEQQGSALEH HSQCSADVKR FNSANEDNVT	QVRTFYTKVL         NEEERKRLCE         NIANHLKDAQ         LFIQRKAVKN         480         981 50         834 44         720 39         623 34         476 27         361 24         248 1	16 147 1 <sup>4</sup>
481 FTDVHPDYGA RVQALLDQYN SQKPKNAIHT YVQAGSHIAA	<u>KGKANL</u> 526 1805.93 1691.89 1505.81 1390.78 1277.70 1178.63 1121.0	61 1007.57
The matched peptides cover 12% (64/526 AA's) of the protein.		

Peptide Selector - Agilent Spectrum Mill Rev. 3.3.078



#### **Catalase Peptide EAE – Peptide Selector**





### **External Calibration on Catalase Peptide** Linearity : five order of magnitude





from 78 amol to 7800 fmol







#### **Catalase Quantitation Results**

	Sample					EAE Met	EAE Results							Qua	lifier (82	8.1 -> 720.1) Results	EAE* (ISTD) Results		Qualifier (831.8 -> 727.3) Results			
(	0	Ÿ	Name	Туре	Level	Acq. Date-Time	Exp. Conc.	RT	Resp.	S/N	М	Calc. Conc.	Final Conc.	Accuracy	Ratio	S/N	MI	RT	Resp.	Ratio	S/N	МІ
Þ		٣	RL_Veh_1ppb	Cal	1	2/27/2008 7:49 AM		32.226	85633	169.81	Γ	14.6295	14.6295		10.9	44.30		32.145	5853	20.6	2.71	
			RL_HDose_1ppb	Cal	1	2/27/2008 3:09 PM		31.611	104655	298.12	Γ	54.3647	54.3647		10.6	71.79		31.611	1925	11.2	1.03	

Sample					LAQ Met	LAQ Results							Qualif	ier (638	5 -> 964.2) Results	LAQ* (ISTD) Results		Qualifier (642.2 -> 964.2) Results		.2 -> 964.2) Results	
٢	Ÿ	Name	Туре	Level	Acq. Date-Time	Exp. Conc.	RT	Resp.	S/N	MI	Calc. Conc.	Final Conc.	Accuracy	Ratio	S/N	М	RT	Resp.	Ratio	S/N	MI
▶		RL_Veh_1ppb	Cal	[1	2/27/2008 7:49 AM		16.291	4781	71.89	Γ	5.8898	5.8898		55.4	35.63		16.268	812	49.9	13.65	
		RL_HDose_1ppb	Cal	1	2/27/2008 3:09 PM		15.647	10560	77.02		12.9958	12.9958		51.2	50.83		15.668	813	44.1	6.34	





#### **Catalase Quantitation Results**

Sample	Peptide	Catalase (fmol/µg protein)	Catalase (pg/µg protein)	Fold Change EAETFPFNP FDLTK	Fold Change LAQEDPDYG LR	Fold Change 2D-DIGE	
Vehicle	EAETFPFNPFDLTK	8.84	14.63	1 00	1 00	1.00	
Treated	LAQEDPDYGLR	4.61	5.89	1.00	1.00		
Troglitazone	EAETFPFNPFDLTK	32.69	54.36	27	2.2	1 45	
Treated	LAQEDPDYGLR	10.13	13.00	5.7	2.2	1.45	



### HPLC-Chip/QQQ System For Biomarker Validation

- Offers high sensitivity and large dynamic range
- Provides robust and stable nanoflow with HPLC-Chip
- Demonstrates good retention time and MS reproducibility
- Peptide Selector helps determine SRM transitions



#### **Acknowledgements**

#### University College Dublin, Conway Institute

**Stephen Pennington** 

**Ben Collins** 

Thomas Lau

William M. Gallagher

#### Novonordisk

Albrecht Gruhler

#### Sanofi-Aventis

Jean-Charles Gautier

#### **Agilent Technologies**

Ning Tang Peter Stone



#### www.proteomics-lab.com



#### Proteomics



Adilent's protein HPCL-Chin incorporate



### **Wrap-up E-Seminar Questions**

## Thank you for attending Agilent e-Seminars. Our e-Seminar schedule is expanding every week. Please check our website frequently at www.agilent.com/chem/education

#### Or register for



to receive regular updates

