

PRISCILA IAMASHITA

**Estudo dinâmico da expressão gênica global durante a interação STEC-enterócito utilizando séries temporais**

Tese apresentada à Faculdade de Medicina da  
Universidade de São Paulo para obtenção do  
título de Doutor em Ciências

Programa de Pediatria

Orientador: Prof. Dr. Carlos Alberto Moreira-Filho

**São Paulo**

**2017**

PRISCILA IAMASHITA

**Estudo dinâmico da expressão gênica global durante a interação STEC-enterócito utilizando séries temporais**

Tese apresentada à Faculdade de Medicina da  
Universidade de São Paulo para obtenção do  
título de Doutor em Ciências

Programa de Pediatria

Orientador: Prof. Dr. Carlos Alberto Moreira-Filho

**São Paulo**

**2017**

**Dados Internacionais de Catalogação na Publicação (CIP)**

Preparada pela Biblioteca da  
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Iamashita, Priscila  
Estudo dinâmico da expressão gênica global  
durante a interação STEC-enterócito utilizando séries  
temporais / Priscila Iamashita. -- São Paulo, 2017.  
Tese(doutorado)--Faculdade de Medicina da  
Universidade de São Paulo.  
Programa de Pediatria.  
Orientador: Carlos Alberto Moreira Filho.

Descritores: 1.Escherichia coli Shiga toxigênica  
2.Síndrome hemolítico-urêmica 3.Células Caco-2  
4.Interações hospedeiro-patógeno 5.Biologia de  
sistemas 6.Redes reguladoras de genes

USP/FM/DBD-414/17

*Aos meus pais,  
por terem me dado toda a base necessária  
para que eu alcançasse as minhas metas  
e me tornasse quem eu sou hoje.*



## AGRADECIMENTOS

Ao meu orientador Prof. Dr. Carlos Alberto Moreira-Filho a quem sou muito grata pelos ensinamentos e pelo meu desenvolvimento profissional.

À minha coorientadora Dra. Silvia Yumi Bando, em especial, pela imprescindível ajuda neste trabalho, pela contribuição para o meu crescimento profissional e pela amizade.

À especialista em pesquisa do Laboratório de Genômica Funcional do Departamento de Pediatria da FMUSP, Dra. Fernanda Bernardi Bertonha por toda ajuda, pela revisão deste trabalho, e pela amizade.

Ao técnico do LIM 36, Leandro Rodrigues Ferreira por ser sempre colaborativo em nosso trabalho experimental.

Aos meus colegas de Doutorado, Hatylas Azevedo e Nathália Khaled pelo companheirismo de laboratório.

Aos nossos colaboradores Prof. Dr. Luciano da Fontoura Costa e Dr. Filipi Nascimento Silva do Instituto de Física de São Carlos, IFSC-USP pela preciosa colaboração neste trabalho e grande paciência em nos ensinar programação.

Ao colaborador Prof. Dr. André Fujita do Instituto de Matemática e Estatística, IME-USP pelas sugestões, pelos scripts utilizados no início das análises e pelas discussões produtivas.

À colaboradora Profa. Dra. Beatriz E. C. Guth do Departamento de Microbiologia, Imunologia e Parasitologia da UNIFESP pelas cepas de bactérias gentilmente cedidas.

À colaboradora Dra. Cecília M. Abe do Laboratório de Biologia Celular do Instituto Butantan pelas imagens de Microscopia Eletrônica de Varredura (MEV).

À técnica do LIM 36, Fernanda Macaferri por toda ajuda técnica.

À secretária do professor Carlos, Regina Célia de Paula Pena pela colaboração no andamento deste trabalho.

À todos os funcionários e alunos do LIM 36 que de uma forma ou de outra contribuíram para o andamento da tese.

Ao meu marido, Bruno Eizo Higaki por todo incentivo, apoio, paciência, compreensão, pela companhia aos finais de semana no cultivo de células e pelas suas dicas.

À minha família querida: meu pai Alberto Massayoshi Iamashita (*in memoriam*), minha mãe Kayoko Kawakami Iamashita, meus irmãos: Edy Iamashita e Zuleika Iamashita, e meu cunhado Marcio Henrique Nakamura meus agradecimentos por estarem sempre ao meu lado me dando todo o suporte e apoio necessários para a concretização desta e de outras etapas da minha vida.

Aos meus sogros, Elisa Higaki e Jorge Higaki, cunhado Ivan Ryuzo Higaki e namorada Lyandra Maki Nakagawa, por todo incentivo e torcida.

À todos os meus amigos e familiares pela compreensão e pela torcida.

À Universidade de São Paulo (USP) e ao Instituto da Criança (ICr) por me receberem como aluna de Pós-Graduação.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudos de Doutorado.

*“No mundo são necessários no máximo seis laços de amizade  
para que duas pessoas quaisquer estejam ligadas.”*

Stanley Milgran, 1967

## SUMÁRIO

Lista de figuras	
Lista de tabelas	
Lista de anexos	
Resumo	
Abstract	
1 INTRODUÇÃO.....	1
1.1 <i>Escherichia coli</i> .....	1
1.2 <i>Escherichia coli</i> produtora de toxina Shiga (STEC) .....	1
1.2.1 Fatores de virulência.....	2
1.2.2 Síndrome Hemolítica Urêmica (SHU).....	3
1.2.3 Epidemiologia.....	4
1.3 Abordagem de biologia de sistemas.....	5
1.4 Rede de coexpressão gênica (RCG).....	6
1.4.1 Conectividade.....	7
1.4.2 Detecção de comunidades.....	7
1.5 Construção e análise de RCG.....	9
1.6 Estudo de séries temporais.....	11
1.6.1 RCG após interação STEC-enterócito.....	13
2 OBJETIVOS.....	15
2.1 Objetivo geral.....	15
2.2 Objetivos específicos.....	15
3 MATERIAIS E MÉTODOS.....	16
3.1 Amostras bacterianas.....	17
3.2 Cultura celular e ensaios de interação temporal STEC-enterócito....	18
3.3 Extração de RNA das células Caco-2.....	18
3.4 Hibridização dos microarranjos de DNA.....	19
3.5 Análise das redes de coexpressão gênica (RCG).....	20
3.5.1 Série temporal.....	20
3.5.2 Análise topológica das redes.....	20
3.5.2.1 Obtenção da matriz de expressão gênica.....	20
3.5.2.2 Análise comparativa e construção de rede de coexpressão gênica.....	20
3.5.3 <i>Weighted gene coexpression network analysis</i> (WGCNA).....	22
3.5.3.1 Obtenção da matriz de expressão gênica.....	22

3.5.3.2 Construção de rede e detecção de módulos transcripcionais.....	22
3.5.3.3 Correlação dos módulos transcripcionais com cada janela (JJ).....	24
3.5.3.4 Detecção de hubs dos módulos transcripcionais.....	24
3.6 Microscopia eletrônica de varredura (MEV).....	25
3.7 Validação dos microarranjos por RT-qPCR.....	25
4 RESULTADOS.....	28
4.1 Qualidade da hibridização e obtenção dos dados de expressão.....	28
4.2 Análise da variação topológica de redes de coexpressão gênica.....	31
4.2.1 Redes de Caco2/EH41.....	31
4.2.2 Redes de Caco-2/Ec472.....	33
4.3 WGCNA.....	33
4.3.1 Redes e identificação de módulos transcripcionais.....	33
4.3.2 Caracterização funcional dos módulos transcripcionais.....	36
4.3.2.1 Rede Caco-2/EH41.....	36
4.3.2.2 Rede Caco-2/Ec472.....	49
4.4 MEV.....	62
4.5 Validação técnica dos microarranjos de DNA.....	64
5 DISCUSSÃO.....	73
5.1 Caco-2/EH41.....	76
5.2 Caco-2/Ec472.....	80
6 CONCLUSÕES.....	85
7 ANEXOS.....	87
8 REFERÊNCIAS.....	117

## LISTA DE FIGURAS

Figura 1 - Distribuição dos nós pelo teste Kolmogorov-Smirnov para a verificação do status <i>scale free</i> .....	7
Figura 2 - Representação de uma rede com estrutura de comunidades.....	8
Figura 3 - Representação esquemática das alterações das propriedades das redes na transição da saúde para a doença.....	13
Figura 4 - Etapas envolvidas na análise dos dados de expressão gênica global dos microarranjos de DNA em estudos de redes de interação transcricional (A a D).....	17
Figura 5 - Etapas envolvidas para a construção das RCG para cada JS (A a E) .....	21
Figura 6 - Etapas envolvidas para a construção e análise de RCG bem como a análise dos módulos transcricionais associados às janelas JJ (A a E).....	24
Figura 7 - Gráfico boxplot dos valores de expressão dos microarranjos de DNA das amostras T0 a T24 de Caco-2/EH41.....	29
Figura 8 - Gráfico boxplot dos valores de expressão dos microarranjos de DNA das amostras T0 a T24 de Caco-2/Ec472.....	30
Figura 9 - Redes de coexpressão gênica dos grupos Caco-2/EH41 ou Caco-2/Ec472 obtidas das janelas JS1 a JS10. Teste Kolmogorov-.Smirnov (K-S) para avaliar a topologia das redes (status <i>scale-free</i> ) para as redes do grupo Caco-2/EH41 (A) e do grupo Caco-2/Ec472(D).....	32
Figura 10 - Dendrograma das amostras para exclusão de <i>outliers</i> .....	34
Figura 11 - Curva da variação topológica da rede para diferentes valores $\beta$ ....	34
Figura 12 - Dendrograma dos genes e módulos transcricionais.....	35
Figura 13 - Rede eigengene dos módulos e das janelas do grupo Caco2/EH41 e matriz de correlação entre os módulos e janelas....	37
Figura 14 - Associação módulo-janela do grupo Caco-2/EH41.....	39
Figura 15 - Histograma mostrando as principais vias metabólicas e número de genes dos módulos associados às janelas da rede do grupo Caco-2/EH41.....	40
Figura 16 - Gráfico de dispersão dos valores de kTotal (kT) vs kWithin (kIn) dos genes dos módulos associados às janelas da rede do grupo Caco-2/EH41.....	42

Figura 17 - Análise funcional dos hubs dos módulos associados às janelas da rede do grupo Caco-2/EH41.....	43
Figura 18 - Rede eigengene dos módulos e das janelas de Caco-2/Ec472 e matriz de correlação entre os módulos e janelas..	50
Figura 19 - Associação módulo-janela do grupo Caco-2/Ec472.....	52
Figura 20 - Histograma mostrando as principais vias metabólicas e número de genes dos módulos associados às janelas da rede do grupo Caco-2/Ec472.....	53
Figura 21 - Gráfico de dispersão dos valores de kTotal (kT) vs kWithin (kIn) dos genes dos módulos associados às janelas da rede do grupo Caco-2/Ec472.....	55
Figura 22 - Análise funcional dos hubs dos módulos associados às janelas da rede do grupo Caco-2/Ec472.....	56
Figura 23 - Imagens das células Caco-2 em diferentes intervalos de tempo (T0, T5, T15 e T24) durante a interação com EH41 ou Ec472/01 obtidas por MEV.....	63
Figura 24 - Gráfico de dispersão dos genes em cada módulo da rede do grupo Caco-2/EH41.....	65
Figura 25 - Gráfico de dispersão dos genes em cada módulo da rede do grupo Caco-2/Ec472.....	66
Figura 26 - Mediana dos valores de expressão obtidos por microarranjos de DNA dos genes endógenos candidatos para a escolha de genes de referência para o grupo Caco-2/EH41.....	68
Figura 27 - Mediana dos valores de expressão obtidos por microarranjos de DNA dos genes endógenos candidatos para a escolha de genes de referência para o grupo Caco-2/Ec472.....	69
Figura 28 - Curva de amplificação dos genes endógenos candidatos a genes de referência por qPCR.....	70
Figura 29 - Validação dos dados de expressão obtidos por microarranjos de DNA do grupo Caco-2/EH41.....	71
Figura 30 - Validação dos dados de expressão obtidos por microarranjos de DNA do grupo Caco-2/Ec472.....	72

## LISTA DE TABELAS

Tabela 1 - Lista dos primers utilizados para a validação técnica dos dados obtidos por microarranjos de DNA para os grupos Caco-2/EH41 e Caco-2/Ec472.....	27
Tabela 2 - Distribuição do número de genes nos 11 módulos das redes de Caco/EH41 e de Caco-2/Ec472.....	36
Tabela 3 - Descrição da função biológica dos hubs da rede do grupo Caco-2/EH41.....	44
Tabela 4 - Descrição da função biológica dos hubs da rede do grupo Caco-2/Ec472.....	57



## LISTA DE ANEXOS

Anexo 1 - Análise de qualidade da hibridização nos microarranjos de DNA de cada amostra do grupo Caco-2/EH41.....	87
Anexo 2 - Análise de qualidade da hibridização nos microarranjos de DNA de cada amostra do grupo Caco-2/Ec472.....	87
Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41	88
Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472	100

## RESUMO

Iamashita P. *Estudo dinâmico da expressão gênica global durante a interação STEC-enterócito utilizando séries temporais* [tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2017.

As *Escherichia coli* produtoras da toxina *Shiga* (STEC) são importantes patógenos humanos, causando desde diarreias até a síndrome hemolítica urêmica (SHU). Há diversos sorotipos associados a SHU, tais como O157:H7 e O113:H21. No Brasil o sorotipo O113:H21 ainda não aparece associado a SHU, embora seja frequentemente isolado de carcaças e fezes bovinas. Nosso grupo já investigou comparativamente as redes de coexpressão gênica (RCG) de STEC EH41 (associado à SHU) e Ec472/01 (isolado de fezes bovinas). A análise comparativa do perfil transcricional de EH41 e Ec472/01 revelou que somente EH41 expressa um conjunto de genes que inclui o regulador transcricional *dicA*. A maioria destes genes está situada em um único módulo transcricional e podem estar associados a fatores de virulência. Assim, este trabalho centrou-se numa abordagem de biologia de sistemas, integrando análises genômica e fenotípica da resposta de enterócitos Caco-2 à EH41 e Ec472/01. A análise genômica baseou-se no estudo temporal de RCG para compreender os mecanismos moleculares envolvidos na patogenicidade desses dois isolados. As alterações fenotípicas ocorridas nas células Caco-2 ao longo da exposição a cada um dos isolados de STEC foram visualizadas através de MEV. A análise genômica mostrou que o mecanismo molecular da resposta de Caco-2 durante a interação com EH41 ou Ec472/01 é claramente distinto. Nas redes do grupo Caco-2/EH41 as alterações topológicas incluíram a perda do status *scale free* e a sua recuperação, com o estabelecimento de uma nova hierarquia de genes na rede. Esses resultados se enquadram no modelo de redes para transição saúde-doença: a nova rede representa a resposta adaptativa da célula ao patógeno, o que não significa um retorno à normalidade. Já no grupo Caco-2/Ec472 as redes, após a perda do status *scale*

*free*, não recuperam esse status até o final do período estudado, o que sugere um estado de transição mais prolongado para reorganização da hierarquia da rede. Mais ainda, através da caracterização dos módulos transcricionais, foi possível compreender dinamicamente os mecanismos moleculares envolvidos na resposta diferencial de Caco-2 aos dois isolados aqui estudados. STEC EH41 induz rapidamente a resposta inflamatória e apoptótica a partir da primeira hora de interação enterócito-bactéria. Por outro lado, células Caco-2 em contato com Ec472/01 ativam, a partir de uma hora, a fagocitose e, a partir da segunda hora, expressam moduladores da homeostase imune. A análise fenotípica das células Caco-2 mostrou, de forma nítida, uma maior destruição dos microvilos dos enterócitos em contato com EH41 do que com Ec472/01. Integrando os resultados genômicos e fenotípicos pode-se concluir que EH41 induz em Caco-2 - em comparação com Ec472/01 - maiores e mais rápidas alterações na expressão gênica global, além de uma resposta inflamatória e apoptótica excessiva, levando assim a alterações morfológicas mais pronunciadas nas células Caco-2. Em seu conjunto, esses resultados contribuem para uma melhor compreensão dos mecanismos moleculares envolvidos na patogenicidade das STECs associadas à SHU. Assim, as perspectivas de desenvolvimento deste trabalho deverão incluir a investigação de fatores de virulência e vias moleculares envolvidas na indução das respostas imunes que podem conduzir à SHU.

**Descritores:** *Escherichia coli* Shiga toxigênica; Síndrome hemolítico-urêmica; Células Caco-2; Interações hospedeiro-patógeno; Biologia de sistemas; Redes reguladoras de genes.

## ABSTRACT

Iamashita P. *Dinamic study of global gene expression along STEC-enterocyte interaction using time series* [thesis]. São Paulo: "Faculdade de Medicina, Universidade de São Paulo"; 2017.

Shiga toxin-producing *Escherichia coli* (STEC) O113:H21 strains are associated with human diarrhea and some of these strains may cause hemolytic uremic syndrome (HUS). In Brazil O113:H21 strains are commonly found in cattle but, so far, were not isolated from HUS patients. Previously, our group conducted comparative gene co-expression network (GCN) analyses of two O113:H21 STEC strains: EH41, isolated from a HUS patient in Australia, and Ec472/01, isolated from bovine feces in Brazil. Differential transcriptome profiles for EH41 and Ec472/01 revealed a gene set exclusively expressed in EH41, which includes the *dicA* putative virulence factor regulator. GCN analysis showed that this set of genes constitutes an EH41 specific transcriptional module which may be associated to virulence factors. Therefore, in the present work a system biology approach was conducted to investigate the differential Caco-2 response - genomic and phenotypic - to EH41 (Caco-2/EH41) or to Ec472/01 (Caco-2/Ec472) along enterocyte-bacteria interaction. The genomic analysis was based on temporal GCN data in order to gain a better understanding on the molecular mechanisms underlying the capacity to cause HUS. The phenotypic alterations in Caco-2 during enterocyte-bacteria interaction were assessed by scanning electronic microscopy (SEM). The genomic analysis showed that the molecular mechanism of Caco-2 response to EH41 or to Ec472/01 during enterocyte-bacteria interaction is clearly different. The GCN topological analyses for Caco-2/EH41 group revealed loss of the scale-free status after one hour of interaction, persistence of this condition along the second hour and establishment of a new gene hierarchy thereafter. These events resemble the network mechanism of health-disease transition. The new established network represents an adaptive cell response to the pathogen and not the return to a

“normal” state. Conversely, the networks for Caco-2/Ec472 group showed a slow and progressive loss of the scale-free status without its restoration at the end of the time interval here studied. Through transcriptional module characterization it was possible to reveal the dynamic of the molecular mechanism involved in the Caco-2 differential responses to the STEC isolates. EH41 induces a rapid inflammatory and apoptotic response just after the first hour of enterocyte-bacteria interaction. Instead, the Caco-2 response to Ec472/01 is characterized by phagocytosis activation at the first hour, followed by the expression of immune response modulators after the second hour. SEM phenotypic analysis of Caco-2 cells along enterocyte–bacteria interaction showed more intense microvilli destruction in cells exposed to EH41, when compared to cells exposed to Ec472/01. The integration of genomic and phenotypic data allowed us to conclude that EH41, comparatively to Ec472/01, induces greater and precocious global gene expression alterations in Caco-2, what is related to excessive inflammatory and apoptotic responses. These responses are associated with the pronounced morphological alterations observed by SEM in Caco-2 cells exposed to EH41. Altogether, these results contribute for a better understanding of the molecular mechanism involved in STEC pathogenicity associated to HUS. Therefore, the future perspectives for the development of the present work should include the investigation of virulence factors and molecular pathways involved in the induction of immune responses leading to HUS.

**Descriptors:** Shiga-toxigenic *Escherichia coli*; Hemolytic-uremic syndrome; Caco-2 cells; Host-pathogen interactions; Systems biology; Gene regulatory networks.

# 1 INTRODUÇÃO

As doenças diarreicas estão entre as principais causas de mortalidade entre crianças menores de cinco anos de idade, segundo a *World Health Organization* (WHO, 2016): cerca de 840.000 mortes de crianças por essa causa ocorrem em todo o mundo a cada ano, principalmente em países em desenvolvimento. *Escherichia coli* é o agente infeccioso bacteriano mais frequente nas formas endêmicas da diarreia infantil no mundo (WHO, 2012).

## 1.1 *Escherichia coli*

A bactéria *Escherichia coli* (*E. coli*) é uma bacilo Gram-negativo encontrado na microbiota normal do intestino humano e de alguns animais. Esse grande e diverso grupo de bactérias pode ser encontrado em ambientes terrestres e aquáticos. Embora a maioria das espécies seja inofensiva, algumas podem causar sérias doenças. As cepas reconhecidamente patogênicas dessa bactéria provocam infecções no trato urinário, sepse/meningite e doenças entéricas/diarreicas (Kaper et al., 2004).

As principais categorias ou patótipos de *E. coli* diarreiogênicas, classificadas de acordo com as características de virulência, interação com a mucosa intestinal e quadro clínico são: *E. coli* enteropatogênica (EPEC), *E. coli* enterotoxigênica (ETEC), *E. coli* enteroinvasora (EIEC), *E. coli* enteroagregativa (EAEC), *E. coli* difusamente aderente (DAEC) e *E. coli* produtora de toxina Shiga (STEC) (Nataro; Kaper, 1998). Os patótipos de *E. coli* são caracterizados pelo antígeno O determinado pela porção de lipopolissacarídeo da parede celular (LPS) e antígeno H da proteína flagelar que definem o sorogrupo (antígeno O) e o sorotipo (antígenos O e H). São conhecidos em torno de 187 antígenos O e 53 antígenos H que podem formar diversas combinações (Liu et al., 2015; Fratamico et al., 2016).

## 1.2 *Escherichia coli* produtora da toxina Shiga (STEC)

Todas as *E. coli* que possuem o gene *stx* são classificadas como STEC, isto é *E. coli* produtoras da toxina Shiga (dos Santos et al., 2010). Atualmente são conhecidos mais de 400 sorotipos O:H de *E. coli* que possuem o gene *stx*,

sendo o sorotipo O157:H7 o mais importante como causador de infecções em seres humanos (Blanco et al., 2004). Dentre os sorogrupos não-O157, o sorotipo O113:H21 também é conhecido por causar infecções intestinais e extra-intestinais, sendo encontrado em diferentes regiões geográficas, como Europa e Oceania (Karmali et al., 1985; Paton et al., 1999; Monaghan et al., 2012; Bielaszewska et al., 2006; Feng et al., 2014).

### 1.2.1 Fatores de virulência

As complicações clínicas causadas por STEC estão associadas à produção da toxina Shiga (Stx), codificada pelo gene *stx* (Karmali et al., 1985). São conhecidos dois tipos de Stx: Stx1 e Stx2, que formam os subtipos: Stx1a, Stx1c, Stx1d e Stx2a, Stx2b, Stx2c, Stx2d, Stx2e, Stx2f, Stx2g, respectivamente (Scheutz et al., 2012). A toxina Stx2 é mais prevalente em SHU e colite hemorrágica do que Stx1 (Croxen; Finlay, 2010).

Algumas cepas de STEC têm a capacidade para produzir lesões na mucosa intestinal do hospedeiro, uma propriedade mediada pela proteína intimina, fator de adesão codificado pelo gene *eae* (*E. coli attaching and effacing*) que está localizado na região genômica LEE ou *Locus for Enterocyte Effacement* (Paton et al., 1998). Estas cepas pertencem a categoria EHEC (*E. coli* enterohemorrágica), consideradas mais virulentas do que aquelas que não carregam esse gene.

Apesar da falta do gene *eae*, o sorotipo O113 pode causar SHU. As cepas *eae* negativas possuem outros fatores de aderência e virulência associados, tais como Saa, STEC *agglutinating adhesin* (Paton et al., 2001) e SubAB, *subtilase cytotoxin* (Paton et al., 2004), ambos localizados no megaplasmídeo pO113 (Funk et al., 2013). Paton et al. (2001) mutagenizaram o gene *saa* do megaplasmídeo da O113:H21, o que resultou na redução de aderência da bactéria. De acordo com Paton e Paton (2010), essa cepa, envolvida num surto ocorrido em 1998 na Austrália (Paton et al., 2001), expressava SubAB com propriedades citotóxicas e apoptóticas. A proteína autotransportadora da membrana externa, denominada Sab, além de promover a aderência em células epiteliais humanas, contribui para a formação de biofilme (Herold et al., 2014). Muitas cepas também produzem enterohemolisina, uma proteína

codificada pelo gene *ehxA*, porém o seu papel na patogênese de STEC permanece incerto (Feng; Reddy, 2013). Outra proteína, codificada pelo gene *lpf* (*long polar fimbria*), pode estar envolvida no mecanismo de adesão identificado inicialmente em STEC O113:H21 e, posteriormente, em outros sorogrupos de STEC, tanto LEE positivos como negativos. A deleção de *lpf* reduz a virulência da STEC O113:H21 (Doughty et al., 2002).

### **1.2.2 Síndrome Hemolítica Urêmica (SHU)**

*E. coli* produtoras da toxina *Shiga* podem causar desde uma diarreia branda até casos graves de colite hemorrágica que podem evoluir para complicações extraintestinais, como a síndrome hemolítico-urêmica (SHU) e a púrpura trombocitopênica trombótica (PTT), que se caracteriza por uma anemia hemolítica com eritrócitos fragmentados e nucleados, febre, distúrbios neurológicos e disfunção renal (CVE, 2002; Kaper et al., 2004).

Segundo o Centro de Vigilância Epidemiológica (CVE, 2002), pacientes com SHU e PTT são usualmente referidos a unidades de terapia intensiva (UTI) e recebem cuidados especializados que envolvem as áreas de nefrologia, hematologia, neurologia e outras, dependendo das complicações. Sabe-se que 10% dos casos com colite hemorrágica evoluem para SHU, e que cerca de 10 a 50% desses pacientes permanecem com algum comprometimento renal crônico, necessitando de procedimentos dialíticos ou, até mesmo, transplante renal.

Muitos pacientes requerem diálise por longos períodos e cerca de 8% dos casos de SHU apresentam sequelas permanentes como hipertensão, cegueira, paralisias e os efeitos decorrentes da remoção de parte de seus intestinos. Outros 30% continuam com micro-hematúria e graus variáveis de proteinúria, podendo vir a desenvolver insuficiência renal crônica ao longo da vida. A letalidade da SHU varia de 3 e 5% em crianças, podendo ser mais alta em adultos (CVE, 2002).

Complicações neurológicas são mais frequentes na PTT, estando presentes em mais de 90% dos pacientes que evoluíram para óbito. A PTT também apresenta alta letalidade em idosos, onde cerca de 50% dos casos vão a óbito.



No início de 2011 foi relatado um surto de diarreia sanguinolenta causado por STEC do sorotipo O104:H4 com numerosos casos que evoluíram para SHU em pelo menos 13 países da Europa, no Canadá e nos Estados Unidos da América. Somente no período de maio a junho de 2011 já tinham sido notificados pelo menos 4.000 casos, sendo que 900 desses casos evoluíram para SHU e 48 mortes foram constatadas (Askar et al., 2011; Rubino et al., 2011).

### 1.2.3 Epidemiologia

Karmali et al. (1985) foram os primeiros a descrever o sorotipo O113:H21 como produtor da Stx em crianças com SHU. A importância epidemiológica desse sorotipo permaneceu subestimada até o final dos anos 90, quando um surto de SHU em crianças foi relatado na Austrália (Paton et al., 2001). Estudos epidemiológicos realizados no Brasil mostraram, por sua vez, que frequentemente *E. coli* do sorogrupo O113 isoladas de pacientes com diarreia não possuem, ou não expressam, o gene *stx* (Vaz et al., 2004; dos Santos et al., 2010), sendo classificadas como pertencentes às categorias EPEC (*Enteropathogenic E. coli*) ou EAEC (*Enteraggative E. coli*) (dos Santos et al., 2007).

Por outro lado, é importante notar que no Brasil bactérias STEC O113:H21 - que expressam a proteína Stx - são frequentemente isoladas de fezes bovinas (Cerqueira et al., 1997; Irino et al., 2005; dos Santos et al., 2010). Portanto, as bactérias STEC O113:H21 têm potencial para causar doenças como SHU em nosso meio, embora o número de casos relatados seja ainda pequeno (Vaz et al., 2004; dos Santos et al., 2007), ao contrário do que se observa em países como Argentina e Austrália. Nesses países, a contínua presença dessas bactérias nos rebanhos (bovinos e ovinos) tem como contrapartida surtos de diarreia com número importante de casos evoluindo para SHU (Brusa et al., 2013). Em diversos países tem sido relatado aumento da detecção de STEC do sorogrupo O113 e sua associação com casos clínicos de SHU (Majowicz et al., 2014), o que destaca sua importância como patógeno emergente. Apesar disso, este sorotipo ainda é pouco estudado quanto à patogenicidade e epidemiologia (Monaghan et al., 2012). A associação de

STEC com colite hemorrágica, SHU e PTT, e a gravidade dessas doenças, demandam o envolvimento das autoridades de saúde para monitorar não somente o aparecimento de casos e a investigação de surtos, como também estabelecer de forma precisa as fontes de transmissão e aplicar as medidas apropriadas de controle e prevenção (CVE, 2002).

Segundo Feng et al. (2014), que estudaram cepas patogênicas de STEC causadoras de SHU na Austrália, ainda é preciso conduzir mais investigações para esclarecer as diferenças entre essas cepas e as cepas ambientais. Esses autores mostraram que todas as cepas estudadas expressam subtipos de Stx associados com infecções humanas, o que sugere potencial das cepas ambientais para causar a doença. Porém, não foram encontrados padrões genéticos que possam diferenciar uma cepa da outra.

Assim, é fundamental investigar comparativamente o mecanismo celular e molecular das infecções por linhagens de STEC causadoras e não causadoras de SHU no contexto da interação patógeno-hospedeiro. Isso é necessário para se identificar os fatores, bacterianos e do hospedeiro, que determinam o grau de patogenicidade. Uma abordagem experimental para essa investigação é o estudo da interação patógeno-hospedeiro *in vitro*, utilizando-se linhagens celulares que mimetizam o enterócito (p. ex. células Caco-2) e expondo-se essas culturas a linhagens causadoras e não causadoras de SHU. Essa abordagem, associada ao estudo dinâmico (séries temporais) de redes de coexpressão gênica em células expostas ao patógeno, será adotada no presente projeto para a compreensão dos processos de genômica funcional envolvidos na interação STEC-enterócito, como descrito a seguir.

### **1.3 Abordagem de biologia de sistemas**

O avanço da biologia molecular e da bioinformática nesta última década possibilitou a aquisição de uma grande quantidade de informações na área de genômica funcional. A genômica funcional tem por objetivo o estudo global do funcionamento e da interação dos genes e das demais sequências do genoma humano, bem como da influência do meio sobre o funcionamento gênico (epigenética) (Ideker; Krogan, 2012). Campos como a proteômica e

metabolômica também contribuem para o entendimento desses mecanismos moleculares (Carter et al., 2013).

A integração dessas áreas de conhecimento é chamada biologia de sistemas (Chuang et al., 2010; Hillmer, 2015). A biologia de sistemas estuda a integração e o controle da atividade dos genes e de proteínas (Lesk, 2008). Essa abordagem, centrada no estudo global de vias de regulação metabólica (Hillmer, 2015), envolve a análise de grandes volumes de dados: por exemplo, o estudo do transcriptoma (a partir de dados de microarranjos de DNA ou de DNaseq) através de redes de interação gene-gene (Barabási et al., 2011). Ela tem sido utilizada para identificar o mecanismo molecular de diversas doenças (Xu et al., 2010; Bando et al., 2013; Moreira-Filho et al., 2015), bem como biomarcadores funcionais de resposta a terapias (Chu et al., 2008) e de prognóstico (Chuang et al., 2010). Mais ainda, redes de coexpressão gênica (RCG), bem como as redes de proteína-proteína e metabólicas, são governadas por leis universais e as propriedades topológicas e dinâmicas destas redes determinam regras importantes para compreender a organização funcional de células e tecidos (Barabási; Oltvai, 2004; Zhu et al., 2007).

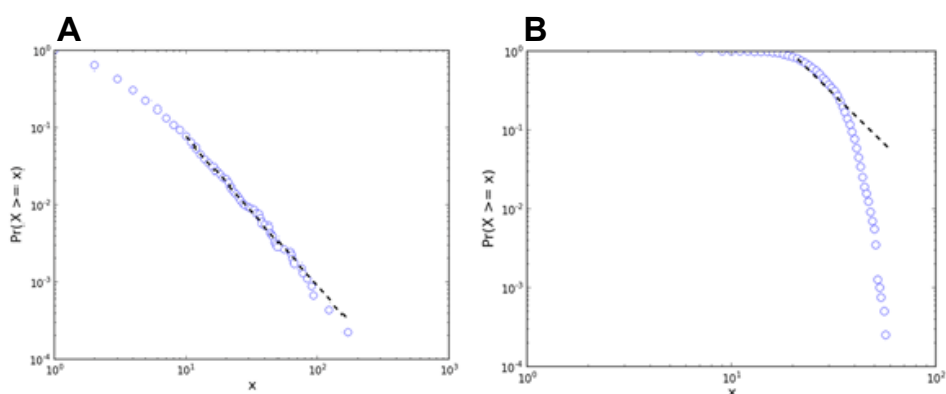
#### **1.4 Rede de Coexpressão Gênica (RCG)**

Uma rede é um grafo representado por um conjunto de vértices ou nós e um conjunto de arestas ou conexões. Cada aresta é especificada por um par de vértices (Lesk, 2008). Redes têm sido muito utilizadas para representar a interação gene-gene no estudo da expressão gênica global (Moreira-Filho et al., 2014a). Uma rede de coexpressão gênica (RCG) mostra a correlação entre os genes. O termo “coexpressão” vem das semelhanças de padrões de expressão dos genes em uma variedade de condições experimentais (Aoki et al., 2008).

As RCG, como as demais redes biológicas, e similarmente às redes sociais e à Internet, não são randômicas e seguem alguns princípios básicos de organização (Newman, 2010). Em redes randômicas (Figura 1B), os nós possuem aproximadamente o mesmo número de ligações (*node degree*) e, portanto, os nós com grande número de ligações são raros. Então, nós em redes randômicas caracteristicamente possuem baixa diversidade de *node*

*degree*. Por outro lado, a maioria das redes “*real world*”, como as RCG ou redes proteína-proteína, são livres de escala (*scale-free*), onde a distribuição de grau segue uma lei de potência. A organização *scale-free* contém *hubs* com muitas ligações e um grande número de nós com um número pequeno de ligações (Barabási; Oltvai, 2004). Redes *scale-free* possuem um limitado número de *hubs*, que são frequentemente associados a funções biológicas relevantes e responsáveis pela robustez da rede (Liu et al., 2011; Winterbach et al., 2013).

O teste Kolmogorov-Smirnov (K-S) corresponde à distribuição da lei de potência em dados empíricos (Clauset et al., 2009). A Figura 1 mostra a distribuição de grau em duas redes, sendo que a Figura 1A representa uma rede *scale-free* e a Figura 1B uma rede randômica.



**Figura 1** - Distribuição dos nós pelo teste Kolmogorov-Smirnov para a verificação do status *scale-free*. **A** - Rede *scale-free*; **B** - Rede randômica

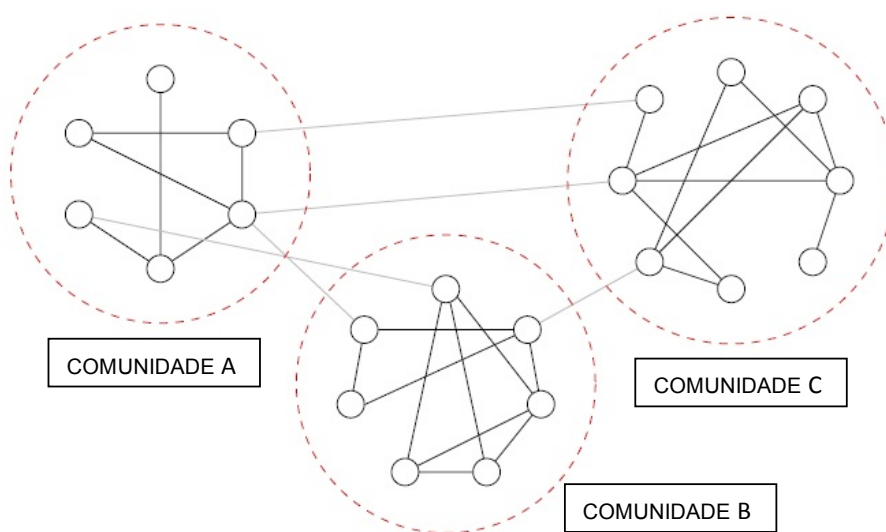
#### 1.4.1 Conectividade

A conectividade é uma medida de correlação que mostra como os componentes de uma rede (nós, vértices) se conectam uns aos outros (Moreira-Filho et al., 2014a,b). A conectividade de uma rede é uma medida importante de sua robustez.

### 1.4.2 Detecção de comunidades

As comunidades podem ser definidas como conjuntos de partições de nós em uma rede que compartilham ligações internas densas e ligações externas esparsas. A detecção de comunidades em redes complexas é obtida encontrando-se a estrutura modular que otimiza a mensuração da modularidade. A modularidade considera a relação entre o número de ligações dentro de uma comunidade e entre nós em distintas comunidades, comparado a um modelo randômico (Newman; Girvan, 2004; Newman, 2010). A Figura 2 ilustra a estrutura de comunidades de uma rede hipotética.

Essa estrutura modular está presente em muitos sistemas reais, tais como redes sociais, redes de conhecimento, internet, etc. E isso também não é diferente para redes de coexpressão gênica. Em RCG a estrutura modular corresponde a um grupo de nós (genes) funcionalmente ligados que atuam em conjunto para realizar uma tarefa específica (Wang; Zheng, 2012; Rosenkrantz et al., 2013).



FONTE: Adaptado de Newman; Girvan, 2004

**Figura 2** - Representação de uma rede com estrutura de comunidades. Cada círculo tracejado representa uma comunidade distinta (A, B ou C), cujas ligações intramódulos estão indicadas em preto e as ligações intermódulos estão indicadas em cinza

## 1.5 Construção e análise de RCG

Redes de coexpressão gênica (RCG) podem ser obtidas para um conjunto de genes, i.e., genes diferencialmente expressos anotados no *Gene Ontology*-GO (rede DE) ou para todos os genes válidos anotados no GO (rede completa ou CO). Essas redes são construídas com base na correlação de covariância gene-gene. Frequentemente são utilizados o coeficiente de correlação de Pearson ou de Spearman, ou medida de similaridade cosseno (Prifti et al., 2008; Song et al., 2012).

Para a construção de RCG são removidas ligações da rede inicialmente completa por meio do aumento gradativo do limiar de correlação (Elo et al., 2007). Após a escolha do limite de força de ligação, geralmente acima de 0,80, a rede é testada quanto ao status *scale-free* (item 1.4) pelo teste de Kolmogorov-Smirnov (K-S), i.e., distribuições de lei de poder em dados empíricos (Clauset et al., 2009).

A visualização e análise das redes ou sub-redes DE podem ser realizadas utilizando-se o software livre Cytoscape (Saito et al., 2012; [www.cytoscape.org](http://www.cytoscape.org)). Por outro lado, a análise da rede CO só é possível através da visualização em 3D. Vários softwares de visualização 3D para redes gene-gene e proteína-proteína estão sendo desenvolvidos (Ishiwata et al., 2009; Pavlopoulos et al., 2008; Wang et al., 2013). Um deles, denominado Networks 3D, desenvolvido pelo grupo de pesquisa do Prof. Dr. Luciano da Fontoura Costa (<http://cyvision.ifsc.usp.br/Cyvision/networks3d>), Instituto de Física de São Carlos, Universidade de São Paulo, é descrito em um artigo do nosso grupo (Bando et al., 2013). Este programa é baseado no algoritmo de força direcionada otimizado que permite a obtenção e visualização das redes em 2D ou 3D (Silva et al., 2013; Bando et al., 2013; Silva et al., 2016).

Utilizando-se o programa Networks 3D em conjunto com o pacote estatístico *Concentric Measurements Software* (<http://cyvision.ifsc.usp.br/concentric/software>) é possível obter medidas relacionadas à hierarquia dos nós e à topologia da rede, tais como a conectividade e a modularidade.

Nosso grupo tem utilizado esta ferramenta para investigar o mecanismo molecular de: i) doenças crônicas não transmissíveis, tais como epilepsia

refratária (Bando et al., 2011; Bando et al., 2013, Moreira-Filho et al., 2015) e câncer (Azevedo et al., 2014); ii) interações patógeno-hospedeiro (Bando et al., 2017); iii) cromossomopatias, p. ex. trissomia do cromossomo 21 e seu impacto no funcionamento do timo (Moreira-Filho et al., 2016). Essa abordagem se revelou muito produtiva: aspectos da topologia e estrutura das redes puderam ser relacionados a mecanismos biológicos relevantes. Por exemplo: i) comunidades em redes de coexpressão gênica correspondem a módulos transcricionais relacionados a determinadas vias metabólicas; ii) alterações na conectividade das redes relacionam-se à transição saúde-doença; iii) perturbações induzidas por agentes ambientais (patógenos, drogas, etc.) levam a alterações topológicas da rede.

Entretanto, esta abordagem não é suficiente para a análise de redes de coexpressão em séries temporais, uma vez que não é possível, em termos práticos, obter o número adequado de amostras e intervalos que permitam correlacionar modificações topológicas das redes, bem como dos módulos transcricionais, ao longo do tempo.

O pacote estatístico *Weighted gene coexpression network analysis* (WGCNA) do software R é uma alternativa para a análise de redes de coexpressão gênica que permite a identificação de módulos transcricionais associados com características fenotípicas ou clínicas distintas. As ferramentas deste pacote incluem a construção de redes, detecção de estruturas modulares, seleção de módulos (*clusters*) de genes, correlação entre os módulos, mensuração da relação entre genes e módulos (*Module Membership*), obtenção de medidas hierárquicas dos nós, cálculos de propriedades topológicas e interação com pacotes de visualização de rede externos (Langfelder; Horvath, 2008).

Esse pacote estatístico é capaz de analisar redes compostas por dezenas de milhares de genes, pois a análise é feita com um número limitado de módulos o que reduz a sua complexidade e permite um maior poder estatístico (Dileo et al., 2011).

Esta metodologia tem sido utilizada com sucesso para a investigação de redes de coexpressão gênica em séries temporais. Liu et al. (2017) utilizaram o WGCNA na identificação de módulos fortemente correlacionados com cada um

dos quatro estágios de diferenciação precoce de cardiomiócitos de células-tronco pluripotentes induzidas por humanos e de células-tronco embrionárias humanas. Este trabalho mostra uma análise temporal de modificações que ocorrem em diferentes estágios de cardiomiócitos.

Hopper et al. (2016) trabalhando com três espécies de videiras diferentes, empregaram o WGCNA na identificação de 30 módulos envolvidos em importantes processos fisiológicos durante a resposta à diferentes tempos de desidratação das folhas. Os autores fizeram o enriquecimento funcional dos módulos e identificaram hubs associados ao metabolismo e sinalização de dois hormônios de estresse vegetal, ácido abscísico e etileno por meio de uma abordagem transcriptômica de séries temporais.

### **1.6 Estudo de séries temporais**

Séries temporais de expressão gênica obtidas a partir de dados de microarranjos de DNA têm sido comumente utilizadas para monitorar a atividade transcricional ao longo de um determinado período de tempo (Whitfield et al., 2002). Essas variações usualmente estão associadas às respostas temporais das células e/ou tecidos a estímulos externos. Esses estímulos podem ser de diferentes naturezas, tais como o crescimento e diferenciação tissular, resposta a um patógeno ou à administração de uma droga (Androulakis et al., 2007; Bar-Joseph et al., 2012).

Dados de expressão de séries temporais permitem a obtenção de informações de grupos de genes que são ativados ou desligados em intervalos de tempo diferentes em resposta a tais perturbações (Bar-Joseph et al., 2012).

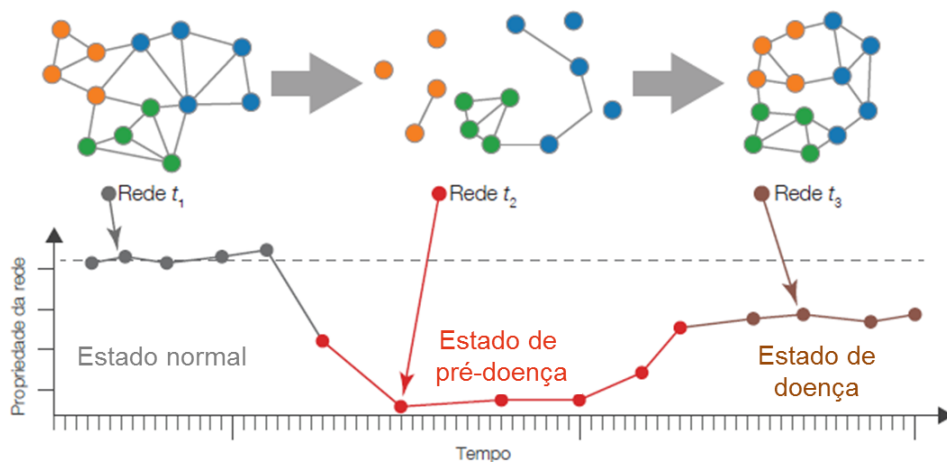
Um dos problemas de processos biológicos dinâmicos é a escolha do ponto de tempo adequado, pois em muitos casos o início do processo estudado é desconhecido, além de poder ser diferente entre os indivíduos. A velocidade do processo pode variar devido a diferenças genéticas entre indivíduos ou à epigenética. Esses problemas podem ser resolvidos pela coleta de múltiplos pontos de tempo para cada amostra. Estatisticamente, o baixo número de amostras é a maior limitação (Elo; Schwikowski, 2013). Outro fator a ser considerado está em como obter material biológico para uma série temporal e



consequentemente, há escassez de dados para desenvolver e/ou aplicar ferramentas matemáticas para a inferência de redes regulatórias.

Perturbações no programa transcricional determinam alterações nas redes de interação gênica, caracterizadas por modificações de conectividade da rede, hierarquia dos genes, composição das comunidades. Essas modificações topológicas nas RCG constituem a força motriz para alterações no metabolismo celular, caracterizando a mudança um estado de equilíbrio para outro (Sieberts; Schadt, 2007; Sahni et al., 2013). O estudo da topologia das RCG permite caracterizar diferentes fenótipos e, muito importante, a transição saúde-doença (Benson; Breitling, 2006; Carter et al., 2013; Bando et al., 2013), como apresentado a seguir.

A transição saúde-doença pode ser esquematicamente dividida em três etapas (Figura 3): um estado normal, um estado de pré-doença (ou estado crítico) e um estado de doença (Liu et al., 2011). No estado normal, as redes gênicas apresentam alta resistência e robustez às perturbações externas. O estado de pré-doença, por sua vez, é caracterizado pela perda de conectividade entre nós e consequente alteração da estrutura de comunidades. A transição entre os estados de saúde e doença pode ocorrer por insulto externo incidindo sobre redes moleculares, determinando redução da robustez da rede. Nesta fase, o sistema é sensível a insultos externos, mas ainda é reversível para o estado normal se ocorrer uma interferência adequada. Nesse estágio, apenas uma pequena mudança nos parâmetros do sistema pode ser suficiente para determinar um colapso transiente da rede (perda de conectividade e, por vezes, subsequente perda do status *scale-free*) com posterior transição para o estado de doença (Liu et al., 2011, Sahni et al., 2013, Sahni et al., 2015).



**Figura 3** - Representação esquemática das alterações das propriedades das redes na transição da saúde para a doença. \*Figura cedida por Filipi Nascimento Silva, IFSC-USP

O estado de doença representa uma nova situação, onde a RCG readquire alta resiliência e robustez, porém com mudanças que podem ser transitórias ou permanentes na hierarquia dos nós e na composição das comunidades. O sistema atinge um novo equilíbrio e, portanto, apresenta dificuldade em retornar ao estado normal. Assim, a identificação das topologias características de cada estado - chamadas de *leading networks* - é fundamental para identificar os mecanismos causais que levam à consolidação do estado de doença (Liu et al., 2011). Nessa linha, uma janela de oportunidade para o estudo dos processos genômicos na transição saúde-doença é a utilização de séries temporais. Conjugada com a utilização de modelos *in vitro*, essa abordagem permite estudar aqueles processos em intervalos sucessivos, igualmente espaçados ou não, com controle rigoroso de variáveis.

### 1.6.1 RCG após interação STEC-enterócito

O nosso grupo já realizou um estudo de interação STEC-enterócito com três horas de interação. Foram analisados o transcriptoma das bactérias e o das células Caco-2. Foram utilizadas duas cepas Stx positivas de STEC O113:H21, sendo uma isolada de paciente com SHU (EH41) e a outra isolada de fezes bovina (Ec472/01), não associada a SHU (Bando et al., 2017).

A análise comparativa das RCG das duas bactérias mostrou que a rede de EH41 expressa um conjunto de genes que não é expresso pela Ec472/01. Mais ainda, a maioria dos genes diferencialmente expressos está hipo regulada em comparação como Ec472/01. Estes genes podem estar envolvidos no crescimento bacteriano, uma vez que EH41 cresce menos que Ec472/01. Este fato pode estar relacionado à adaptação bacteriana para sobreviver por longo período no ambiente intestinal e conseqüentemente causar SHU.

As RCG das células Caco-2 expostas a EH41 ou Ec472/01 mostraram que muitos *hubs* de ambas as redes estão envolvidos na resposta imunológica. Entretanto, a rede Caco-2 exposta à EH41 apresentou maior número de *hubs* associados à inflamação, indicando uma resposta infamatória pronunciada. Além disso, essa rede possui muitas interações negativas e positivas, indicando uma desregulação global da RCG.

Sabe-se que perturbações na rede de interação gênica são a força motriz para alterações no metabolismo celular de um estado de equilíbrio para outro, muitas vezes em resposta a mudanças ambientais (Sieberts; Schadt, 2007; Sahni et al., 2013). Estas transições na topologia das RCG, que refletem alterações na conectividade dos genes, determinam diferentes fenótipos, por exemplo na transição da saúde para a doença (Benson; Breitling, 2006; Carter et al., 2013; Bando et al., 2013).

Por outro lado, o estudo de um único intervalo de tempo não esclarece o mecanismo molecular envolvido na mudança do estado “normal” (rede controle) para o “alterado” (p. ex. rede das células Caco-2 após 3 horas de interação com EH41, Bando et al., 2017). Assim, este trabalho foi centrado na obtenção de redes de interação transcricional a partir de séries temporais de expressão gênica de células Caco-2 expostas a duas linhagens de STEC: uma associada e outra não associada a SHU durante 3 horas de interação já pré-estabelecidos (Bando et al., 2017). Foram também estudadas as alterações temporais na morfologia das células Caco-2 ao longo da interação com STEC. O objetivo principal foi elucidar e comparar os mecanismos moleculares relativos à resposta do enterócito às duas bactérias, como será descrito adiante.

## 2 OBJETIVOS

### 2.1 Objetivo geral

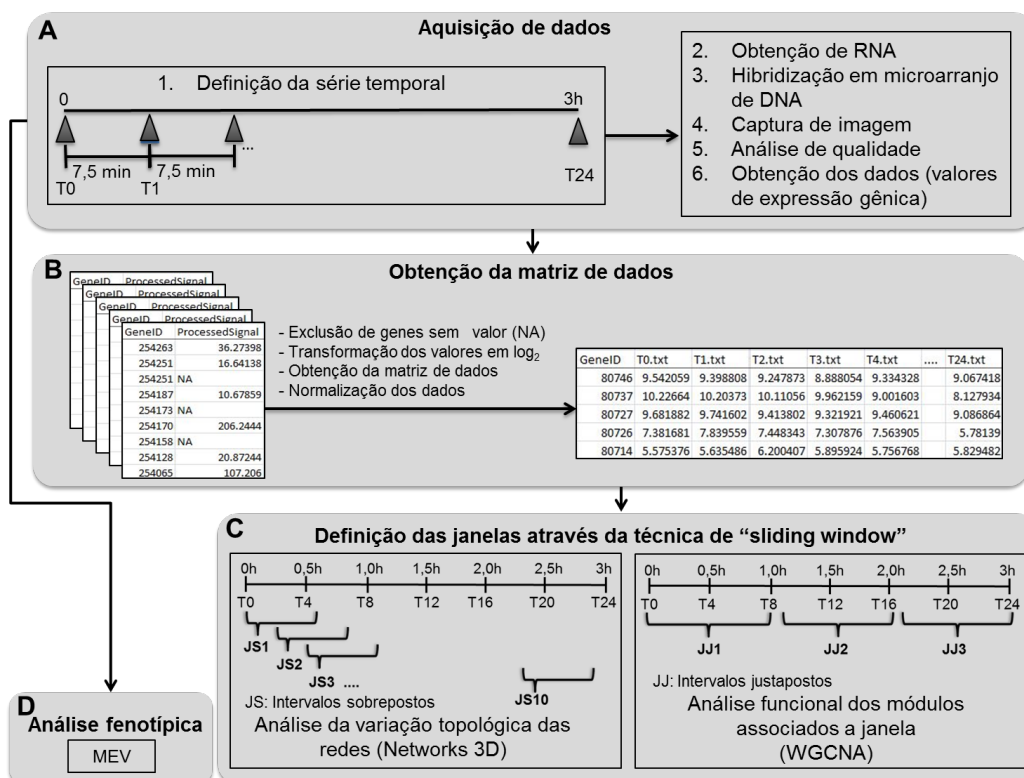
Investigar comparativamente os perfis de expressão gênica global e as alterações morfológicas em séries temporais durante a interação STEC-enterócito *in vitro* para compreender o mecanismo molecular envolvido na resposta das células Caco-2 a duas cepas dessa bactéria, EH41 - isolada de paciente com SHU - e Ec472/01 - isolada de fezes bovina.

### 2.2 Objetivos específicos

- i) Conduzir análises do tipo *sliding-window* para investigar a dinâmica das redes de coexpressão gênica (RCG) ao longo de 3 horas de interação enterócito-STECS;
- ii) Utilizar duas abordagens distintas para a construção de RCG: a) redes baseadas em genes diferencialmente expressos (DE) para análise topológica; b) redes incluindo todos os genes expressos anotados no Gene Ontology (GO) para análise modular;
- iii) Analisar a variação fenotípica temporal das células Caco-2 expostas às STECs utilizando microscopia eletrônica de varredura (MEV).

### 3 MATERIAIS E MÉTODOS

A Figura 4 representa desenho experimental deste trabalho. Inicialmente definiu-se o período de três horas de ensaio de interação *in vitro* de enterócito-STEAC e determinou-se 25 intervalos de tempo igualmente espaçados para a coleta de amostras para análise de genômica funcional ou fenotípica do enterócito (Figura 4A). As redes de coexpressão gênica foram construídas através de duas abordagens (Figura 4C). Uma delas baseou-se na matriz de dados de expressão gênica obtida pela sobreposição de intervalos de tempo e a construção de redes para cada janela a fim de analisar a variação topológica das redes ao longo do tempo. A outra análise baseou-se na matriz obtida pela justaposição de intervalos e a construção de uma única rede para o estudo funcional dos módulos transicionais associados a cada janela. Foram também estudadas as alterações temporais na morfologia da célula Caco-2 ao longo da interação com STEAC (Figura 4D).



**Figura 4** - Etapas envolvidas na análise dos dados de expressão gênica global dos microarranjos de DNA em estudos de redes de interação transcricional (A a D)

### 3.1 Amostras bacterianas

As amostras bacterianas utilizadas pertencem à bacterioteca mantida pelo Departamento de Microbiologia, Imunologia e Parasitologia da Escola Paulista de Medicina da UNIFESP e foram fornecidas pela Profa. Dra. Beatriz E. C. Guth. Dois isolados de STEC O113:H21 foram utilizados: EH41 isolado de uma criança com SHU na Austrália (Karmali et al., 1985) e Ec472/01 isolado de fezes bovina (dos Santos et al., 2010). Uma única colônia de cada amostra foi repicada para o meio Luria Bertani (LB) por 18 horas a 37 °C. Alíquotas destas culturas em fase exponencial foram utilizadas para os ensaios de interação com células Caco-2.

### **3.2 Cultura celular e ensaios de interação temporal STEC-enterócito**

As células Caco-2 foram cultivadas em frascos plásticos estéreis de 25 cm<sup>2</sup> (TPP, Suíça) ou em placas para cultivo celular de quatro poços (Thermo Scientific, EUA) com lamínulas de vidro de 13 mm contendo 4 mL ou 500 mL de Meio Mínimo Essencial de Eagle, modificado por Dulbecco (DMEM; Cultilab, Brasil), respectivamente. O meio de cultura foi suplementado com 10% de Soro Fetal Bovino (SFB; Gibco, EUA) e antibióticos (penicilina 1000 U/mL e estreptomicina 1 mg/mL; Sigma-Aldrich, EUA) sob atmosfera de 5% de CO<sub>2</sub> e incubação à 37 °C. Entre 5 a 7 dias, as células cresceram até atingir 80% de confluência e formar uma monocamada de células epiteliais polarizadas - lado apical apresenta microvilosidades (Halbleib et al., 2007). Vinte e quatro horas antes do ensaio de interação, as células foram lavadas duas vezes com DMEM sem soro e sem antibióticos. Em seguida, foram incubadas com DMEM contendo 5% de SFB e sem antibióticos. Inoculou-se 400 µL ou 50 µL de cultura bacteriana ( $Abs_{550} = 0.35$ ) sobre a monocamada de Caco-2 cultivada em frascos ou em placas, respectivamente. As culturas celulares foram incubadas a 37 °C sob atmosfera de 5% de CO<sub>2</sub> em diferentes intervalos de tempo. Após cada intervalo, as células Caco-2 foram recuperadas para a extração de RNA ou para o procedimento de Microscopia Eletrônica de Varredura - MEV (ver adiante).

### **3.3 Extração de RNA das células Caco-2**

Após cada intervalo de tempo de interação STEC-enterócito, as células foram lavadas por cinco vezes com tampão fosfato salino (PBS; Laborclin, Brasil) 1x e rapidamente adicionou-se 600 µL do tampão RLT de lise celular (Qiagen, Valencia, CA) diretamente no frasco. O lisado foi recolhido em um microtubo, seguindo-se a extração de RNA com o kit RNeasy Mini Kit (cat. no. 74106, Qiagen, Valencia, CA) de acordo com as recomendações do fabricante. A verificação da pureza e a quantificação do RNA foram realizadas no nanoespectrofotômetro NanoVue (GE Healthcare, EUA), através de leitura a 260 e 280 nm. Essas amostras foram armazenadas a -80 °C até a sua utilização. A integridade do RNA foi verificada utilizando a plataforma Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA). Todas amostras

que apresentaram o Número de Integridade do RNA (NIR)  $\geq 7.0$  foram armazenadas a  $-80\text{ }^{\circ}\text{C}$  até o uso nos experimentos de hibridização.

### 3.4 Hibridização dos microarranjos de DNA

Os experimentos de hibridização para análise de expressão gênica foram realizados na plataforma Agilent (Agilent Technologies, Santa Clara, CA). Esta plataforma utiliza *bioarrays* confeccionados com oligonucleotídeos de 60 bases, individualmente pré-validados. Para análises do perfil de expressão gênica foram utilizadas lâminas de microarranjos de DNA contendo 44.000 transcritos de genes humanos (cat. no. G4845A, Agilent Technologies, Santa Clara, CA). Os procedimentos de hibridização utilizando o sistema de uma cor marcada com o fluorocromo *cyanine-3* (Cy3) seguem estritamente os protocolos fornecidos pelo fabricante (*One-Color Microarray-Based Gene Expression Analysis Low Input Quick Amp Labeling* versão 6.5, maio/2010, Agilent, Santa Clara, CA). As imagens foram capturadas pelo scanner Agilent *Bundle* de acordo com os parâmetros recomendados para Agilent *bioarrays*. O software *Feature Extraction* versão 9.5.3 foi utilizado para avaliar o sinal fluorescente da hibridização dos pontos (*spots*), normalizar estes sinais e exportar os dados. A qualidade dos arranjos foi avaliada através da correlação entre a quantidade de *spots* e a intensidade destes utilizando o software R versão 3.3.3 (R Development Core Team, 2014) com scripts feitos *in house*. Um arranjo com boa qualidade mostra distribuição normal de todos os sinais. Por meio da análise de cada arranjo, *spots* com um ou dois *flags* (baixa intensidade, saturação, etc) foram considerados como NA, isto é, sem valor válido de expressão. Ou seja, aqueles transcritos que apresentaram sinais não uniformemente distribuídos com pelo menos um NA foram descartados (Figura 4B). A seguir foi feito a transformação dos valores de expressão em  $\log_2$  e também a normalização de dados de regressão não-paramétrica robusta denominado método Lowess (Limma package, Smyth, 2005; <http://www.bioconductor.org/>).



### **3.5 Análise das redes de coexpressão gênica (RCG)**

#### **3.5.1 Série temporal**

A série temporal adotada foi de intervalos igualmente espaçados de sete minutos e meio ao longo de três horas (Figura 4A). Assim, foram obtidos 25 intervalos de tempo (T0 até T24). Em T0, o RNA das células Caco-2 foram extraídas imediatamente após a inoculação bacteriana (sem período de incubação).

As matrizes de dados de expressão gênica foram obtidas através da técnica de *sliding-window* em séries temporais. Utilizou-se a técnica de sobreposição (Madsen, 2007; Silva et al., 2015) dos intervalos de tempo para a análise topológica de 10 redes ao longo de três horas de interação patógeno-hospedeiro. Enquanto que para a análise funcional dos módulos associados a diferentes intervalos de tempo, utilizou-se a técnica de justaposição (Camargo et al., 2011) dos intervalos de tempo.

#### **3.5.2 Análise topológica das redes**

##### **3.5.2.1 Obtenção da matriz de expressão gênica**

A matriz de dados de expressão gênica foi composta por 10 janelas sobrepostas (JS) como mostra a Figura 4C. Cada janela incluiu dados de expressão gênica de seis intervalos de tempo sequenciais, mas com sobreposição de quatro intervalos: a primeira janela (JS1) foi composta pelos intervalos T0 a T5; a segunda janela (JS2) foi formada por T2 a T7; e assim sucessivamente até JS10 (T18 a T23).

##### **3.5.2.2 Análise comparativa e construção de rede de coexpressão gênica**

A Figura 5 mostra esquematicamente todas as etapas envolvidas na construção das RCG. A comparação dos dados de expressão gênica das janelas JS1 a JS10 foi feita através do uso do programa TMEV (TIGR MultiExperiment Viewer) versão 4.9.0 (Saeed et al., 2003) utilizando a análise de variância (ANOVA) e correção de Bonferroni Ajustado (Figura 5B).

A matriz de dados dos genes diferencialmente expressos (DE) anotados no *Gene Ontology Database* (GO) de cada grupo individualmente (JS1 - JS10) foi

utilizada para a obtenção da matriz de correlação gene-gene por grupo (Figura 5C). A correlação de Pearson foi utilizada para o cálculo da correlação de cada par de genes. O limiar do valor de correlação foi definido após verificar se a topologia da rede JS1 não se alterava dentro de um intervalo de correlação (0,700 a 0,999). A visualização das redes foi realizada utilizando o programa Networks 3D (Figura 5D). O programa foi desenvolvido por Filipi Nascimento Silva do grupo do Prof. Dr. Luciano da Fontoura Costa do Instituto de Física de São Carlos, IFSC-USP (<http://cyvision.ifsc.usp.br/software/networks3d>; Bando et al., 2013; Moreira-Filho et al., 2014a,b; Moreira-Filho et al., 2015). As redes foram analisadas quanto ao status *scale-free* utilizando o teste estatístico Kolmogorov-Smirnov (Figura 5E).

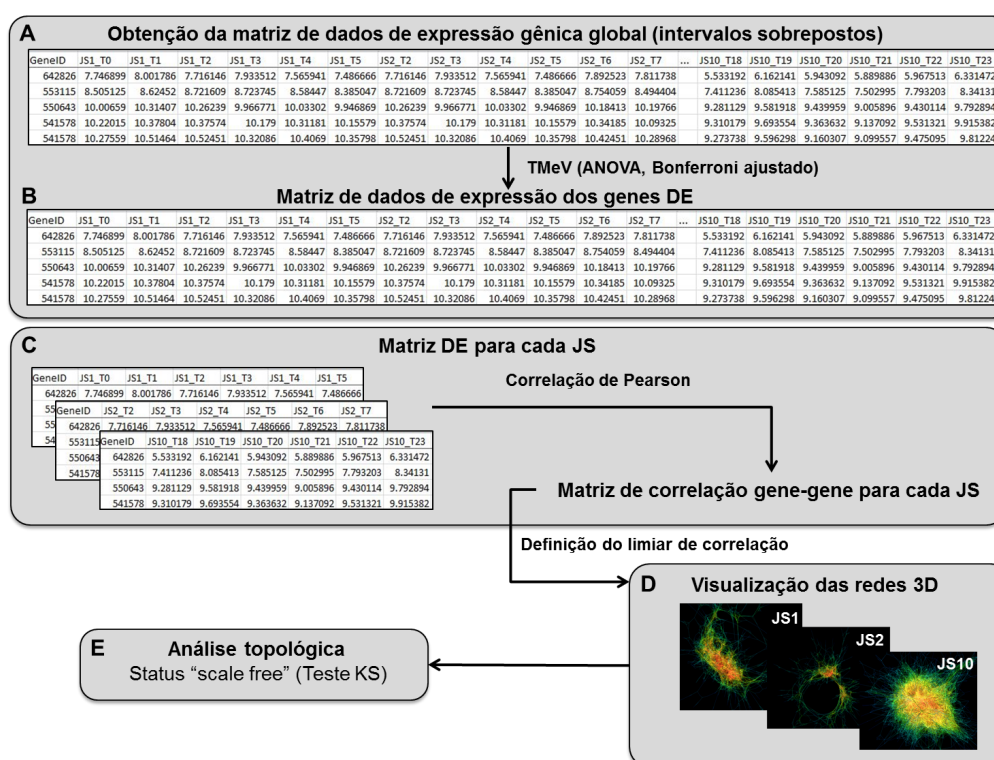


Figura 5 - Etapas envolvidas para a construção das RCG para cada JS (A a E)

### 3.5.3 *Weighted gene coexpression network analysis (WGCNA)*

#### 3.5.3.1 **Obtenção da matriz de expressão gênica**

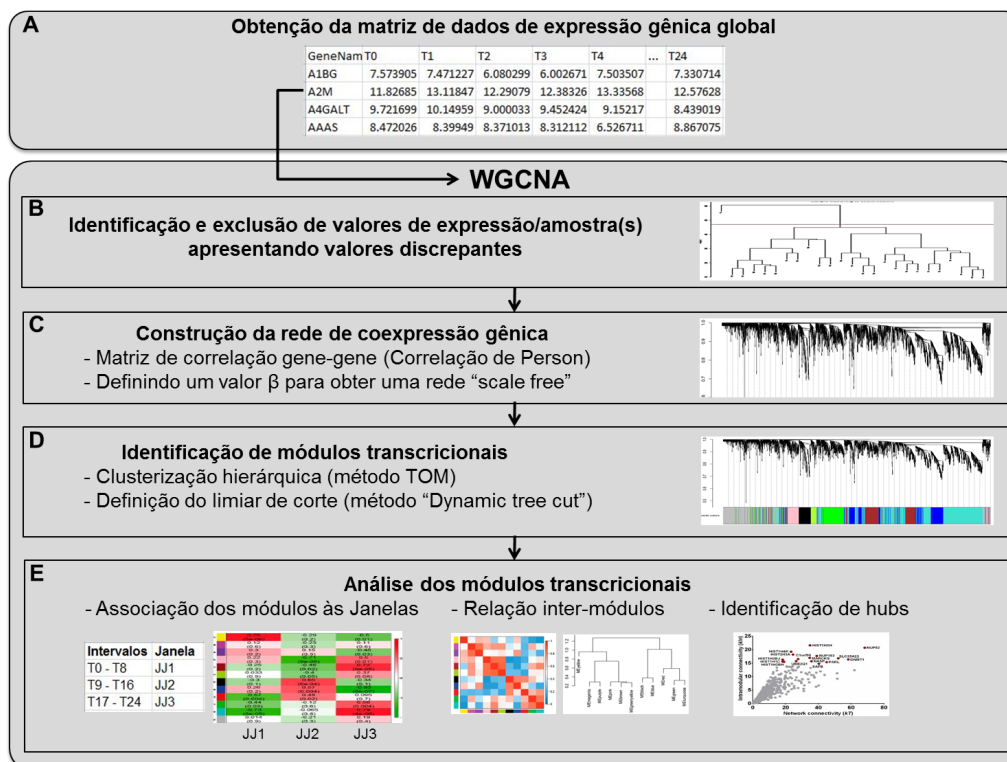
As amostras foram agrupadas em três janelas justapostas (JJ) com nove amostras na primeira janela e oito amostras em cada uma das outras duas janelas. Dessa forma, a janela 1 (JJ1) contém os nove primeiros intervalos de tempo (T0 a T8) da primeira hora de interação, a janela 2 (JJ2) contém os próximos oito intervalos de T9 a T16, e a janela 3 (JJ3) contém os oito últimos intervalos de T17 a T24. A Figura 4C mostra esquematicamente como foi composta cada janela.

#### 3.5.3.2 **Construção de rede e detecção de módulos transcricionais**

As redes de coexpressão gênica global foram construídas utilizando o método *Weighted Gene Coexpression Network Analysis (WGCNA)*, um pacote do software R (Zhang; Horvath, 2005; Langfelder; Horvath, 2008) (Figura 6). WGCNA é utilizado para identificar padrões de correlação entre genes em dados de expressão gênica de microarranjos. Inicialmente os dados de expressão normalizados (item 3.4) das 25 amostras de cada interação Caco-2/EH41 e Caco-2/Ec472 foram checados para a remoção de amostras com valores de expressão discrepantes, ou *outliers* (Figura 6B). As RCG foram construídas utilizando o valor absoluto da correlação de Pearson como medida de similaridade entre os perfis de expressão gênica de todas as amostras. Essa matriz de similaridade foi então transformada em uma matriz de adjacência elevada a uma potência  $\beta$  (*soft-threshold*). Esta etapa enfatiza as correlações fortes e atenua as correlações fracas em uma escala exponencial (mimetizando uma distribuição *scale-free*). As correlações ponderadas ou *weighted* representam as forças de ligação entre os genes na rede (Figura 6C).

A escolha do parâmetro  $\beta$  se baseia no critério de topologia de rede *scale-free* descrito por Zhang e Horvath (2005). Verificou-se visualmente se a topologia *scale-free* foi satisfeita pelo quadrado da correlação entre  $\log(p(k))$  e  $\log(k)$ , onde  $p(k)$  é a distribuição de frequência - modelo do índice de montagem  $R^2$ . Se o  $R^2$  do modelo se aproximar de 1, então existe uma relação de linha reta entre  $\log(p(k))$  e  $\log(k)$ . Quanto maior a potência  $\beta$ , menor é a conectividade média da rede. Geralmente utiliza-se a menor potência  $\beta$  onde a

curva de saturação (relação entre  $R^2$  e  $\beta$ ) é alcançada. O critério para escolha da potência  $\beta$  considera valores em que a rede satisfaça a topologia *scale-free* aproximadamente  $R^2 > 0,80$ . A adjacência do *soft-threshold* foi então transformada em um elemento de matriz de sobreposição topológica (TOM, *topological overlapping matrix*), identificando-se assim as estruturas modulares. Ou seja, foi feita uma conversão da medida de dissimilaridade para agrupamento hierárquico, onde agrupam-se os nós estreitamente conectados formando os módulos da rede. A sobreposição topológica serve como filtro para excluir conexões fracas durante a construção da rede. Pelo agrupamento foi obtido um dendrograma da rede e escolhido um ponto de corte. Os módulos correspondentes aos ramos do dendrograma são estabelecidos pelo *Dynamic Tree Cut* (Figura 6D). Para cada módulo de genes formado foi atribuída uma cor diferente, sendo agrupados no módulo *grey* os genes que não foram atribuídos a nenhum módulo. Para entender como os módulos estão correlacionados entre si, uma rede eigengene foi inferida baseada em um critério de agrupamento hierárquico dos módulos. A rede eigengene foi representada por um dendrograma e um gráfico heatmap (Figura 6E). Módulo eigengene se refere ao primeiro componente principal de um dado módulo, considerado como de significado biológico relevante. O dendrograma eigengene foi formado por meta-módulos correspondentes a grupos de eigengenes correlacionados, representados como ramos (*branches*) no dendrograma. Outra forma de representação é o heatmap, onde módulos altamente correlacionados a outros, ou a um atributo fenotípico, aparecem como quadrados avermelhados, em contraposição aos módulos menos relacionados, que aparecem com quadrados azulados.



**Figura 6** - Etapas envolvidas para a construção e análise de RCG bem como a análise dos módulos transcricionais associados às janelas JJ (**A a E**)

### 3.5.3.3 Correlação dos módulos transcricionais com cada janela (JJ)

Para a correlação dos módulos transcricionais com uma determinada característica de um grupo de amostras - neste estudo corresponde a cada janela de tempo - é necessário obter medidas de significância tanto para o módulo quanto para cada gene. Desta forma, primeiramente foram obtidos os valores de GS (do inglês *Gene Significance*) que é uma medida de correlação entre cada gene e a janela (JJ). A segunda medida é a significância do módulo que é a média de GS dos genes do módulo. Assim, a correlação do módulo com cada JJ foi obtida usando a correlação de Pearson e o valor-p obtido pelo teste t.

### 3.5.3.4 Detecção de hubs dos módulos transcricionais

A identificação dos hubs de cada módulo transcricional apresentando alta associação com pelo menos uma janela foi obtida através da plotagem do tipo *scatter plot* dos valores de kTotal (kT) versus kWithin (kIn) de todos os genes

de um módulo. O valor de  $k_{Total}$  corresponde ao número total de ligações de um gene enquanto que  $k_{Within}$  é o número de ligações de um gene com os outros genes do mesmo módulo.

### 3.6 Microscopia eletrônica de varredura (MEV)

A obtenção das imagens de microscopia eletrônica foi realizada em colaboração com a Dra Cecília M. Abe, do Laboratório de Biologia Celular do Instituto Butantan. Para esses estudos, as monocamadas de células Caco-2 foram preparadas em duplicata nas placas de cultura celular de quatro poços com lamínulas de 13 mm como descrito no item 3.2. Após interação de Caco-2 com STEC, as amostras foram retiradas da estufa em intervalos igualmente espaçados de 37,5 min (T5, T10, T15, T20, T24). Em seguida, foram cuidadosamente lavadas uma vez com 900  $\mu$ L de PBS 1X e mantidas por pelo menos 24 h em solução fixadora Karnovsky a 4 °C. Após fixação, as células foram lavadas com tampão de cacodilato de sódio 1 M por 10 min e pós-fixadas com 1% de tetróxido de ósmio em tampão de cacodilato de sódio 1 M por 30 min. Subsequentemente, as amostras foram lavadas três vezes com água destilada e desidratadas utilizando concentrações crescentes de etanol (50%, 75%, 85%, 95%, 100%). Posteriormente, as preparações foram secadas (método *critical point*) no suporte *stubs* utilizando um sistema de evaporação conhecido com o *sputtering*, onde é feita a deposição de partículas de ouro. Os preparados foram examinados em microscópio eletrônico de varredura (QUANTA 250 - FEI Company, Netherlands) a 12.5 kV.

### 3.7 Validação dos microarranjos por RT-qPCR

Ensaio de RT-qPCR foram realizados para validação dos dados de expressão gênica, selecionando-se três genes para cada grupo: Caco-2/EH41 e Caco-2/Ec472. As sequências dos *primers* para esses genes foram desenhadas utilizando-se o programa *Primer3Plus* (<http://primer3plus.com/>). As sequências dos *primers* estão descritas na Tabela 1. Para obtenção do cDNA, o kit *Superscript III First-Strand Synthesis System* (Invitrogen, Carlsbad, USA) foi utilizado seguindo o manual de instruções do fabricante. As reações de qPCR foram realizadas no termociclador *StepOne Real-Time PCR System* (Applied Biosystems, USA), seguindo as condições de ciclagem: 95°C por 20 s,

40 ciclos de : 95 °C por 3 s, 60 °C por 30 s. Os genes *RPLP0* para o grupo Caco-2/EH41 e *MRPL19* para o grupo Caco-2/Ec472 foram utilizados como genes endógenos de referência. A expressão relativa dos genes foi determinada pelo método da curva padrão (Larionov et al., 2005). Os dados de expressão relativa foram plotados no *GraphPad Prism* 6.0.

**Tabela 1 - Lista dos primers utilizados para a validação técnica dos dados obtidos por microarranjos de DNA para os grupos Caco-2/EH41 e Caco-2/Ec472**

Grupo	Módulo	Gene	Primer forward (5'-3')	Primer reverse (5'-3')	Tamanho (pb)	Referência
<b>Caco-2/EH41</b>	Blue	TBKBP1	CTCTGTTAGCTGGGTGGTG	GTGTGGGAAGGGTTAGCTC	110	NM_014726.2
	Blue	PRIMA1	ATCCAGGGCTCTAGTCTGCA	TCCC TTGCTGTGCCGTATAC	182	NM_178013.3
	Blue	NEGR1	GGACTCCTTTTGTGGCCCT	TCCCCTTCCTCCCTTCCTTC	167	NM_173808.2
	Grey	RPLP0	AAACGAGTCC TGGCC TTGTC	TAICCTCGTCCGACTCCCTCC	196	NM_001002.3
<b>Caco-2/Ec472</b>	Turquoise	TNFRSF1A	TCCAGCTCCACCTATACCCC	GCGGGTTCATCAGTGTCTAG	188	NM_001065.3
	Blue	SLC25A25	CGTGGAGAGAGAGGGAGGAA	CGCCCACCACAAATTCACAG	139	NM_001006641.3
	Blue	BCL7B	CACGCAGCTTTGTTCTCACC	GAGCTGAAGGGGTACTGTGG	123	NM_001197244.1
	Grey	MRPL19	ATCAGGCCCAAAAGCGTAGG	GCAAACACCTGGGTGCGAAG	200	NM_014763.3



## 4 RESULTADOS

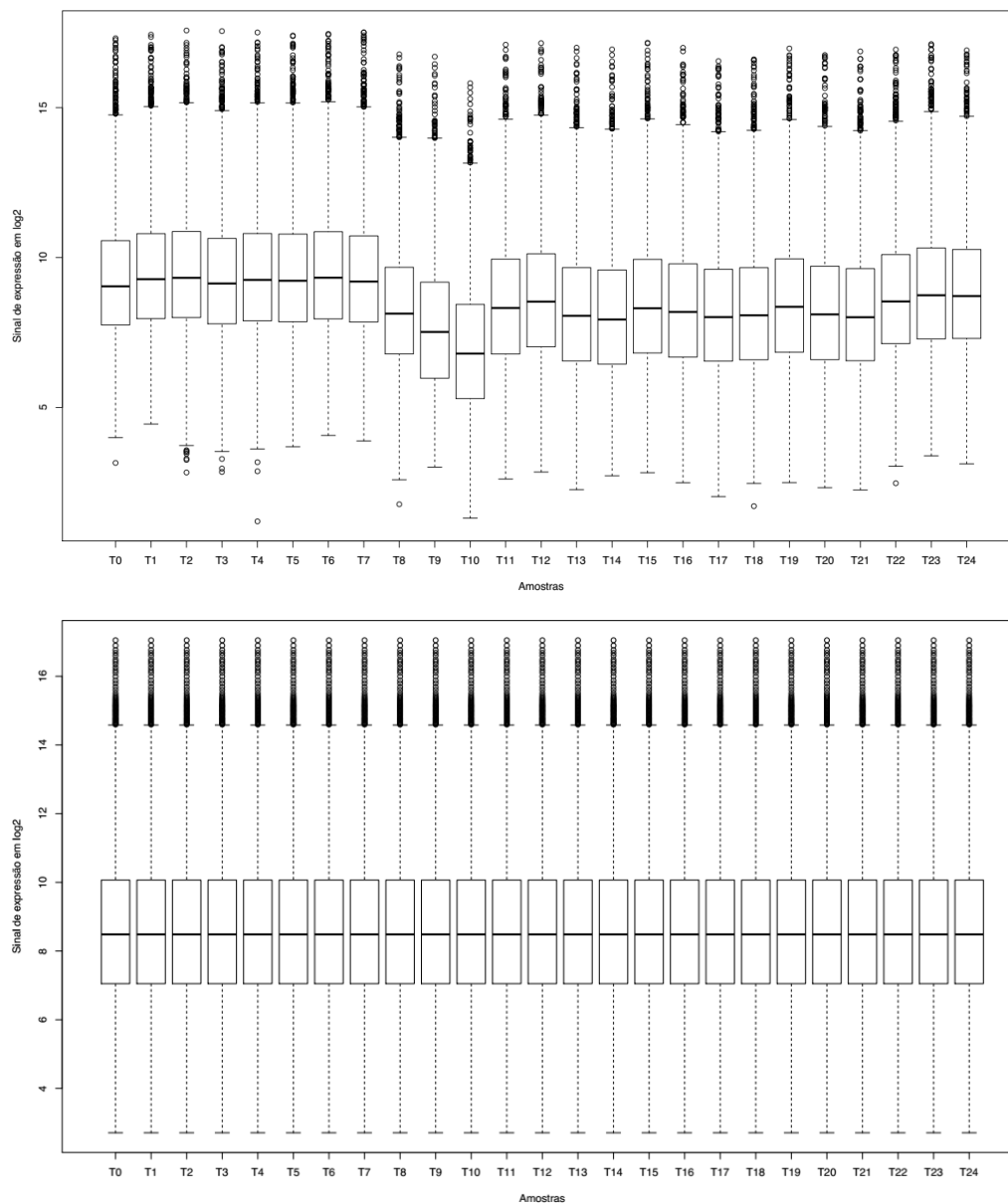
Os resultados deste trabalho compreendem os dados de genômica funcional obtidos em séries temporais e, também, os de MEV que retratam aspectos morfológicos da interação célula-bactéria ao longo de diferentes tempos de exposição das células Caco-2 à EH41 ou Ec472/01.

### 4.1 Qualidade da hibridização e obtenção dos dados de expressão

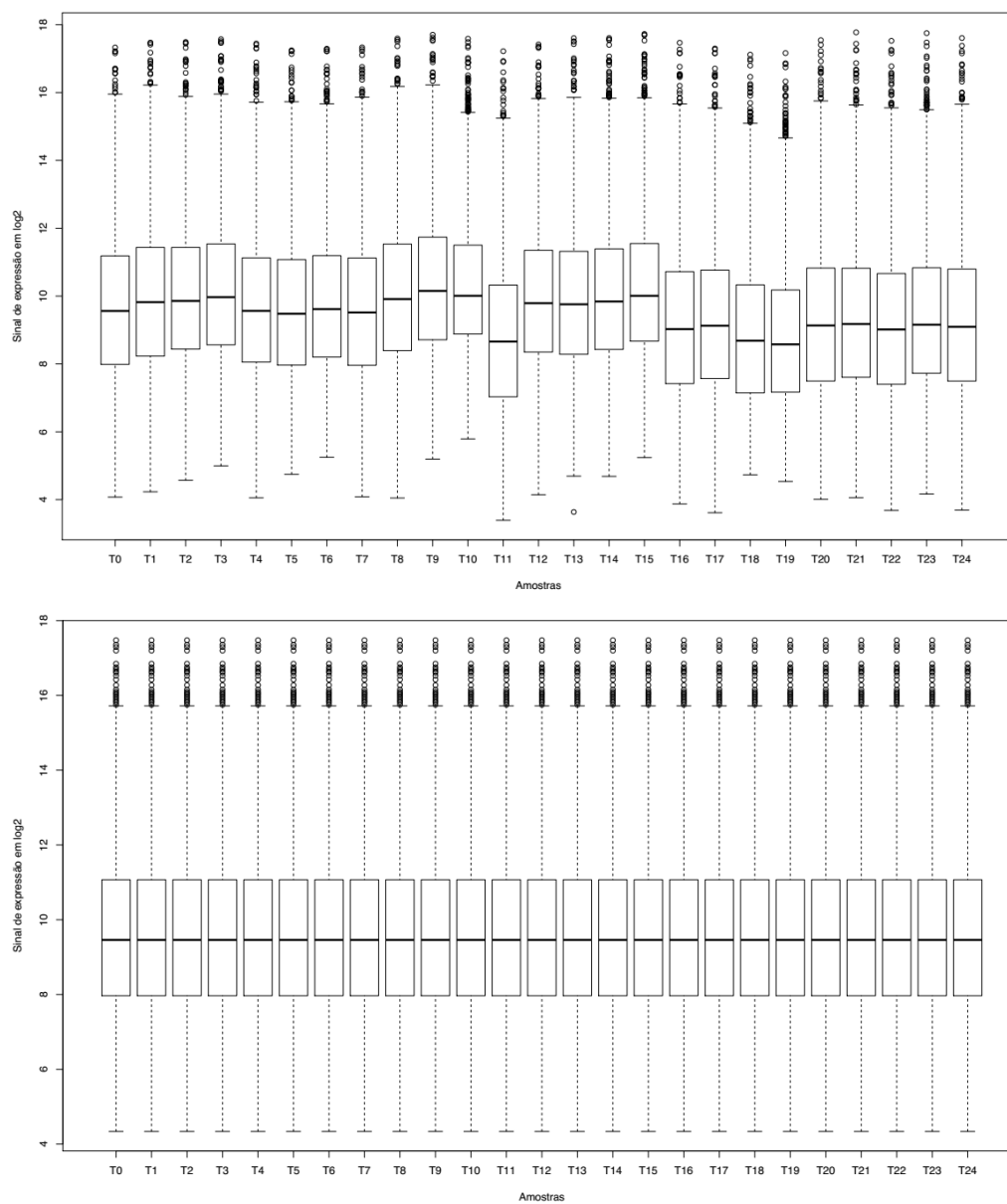
A qualidade da hibridização de todas as amostras do grupo Caco-2/EH41 e do grupo Caco-2/Ec472 foi analisada através da plotagem dos valores de expressão bem como dos valores transformados em  $\log_2$  (Anexos 1 e 2). Todas as amostras apresentaram distribuição normal dos valores de expressão, desta forma todas as amostras foram incluídas no estudo.

A etapa seguinte consistiu na obtenção da matriz de dados de expressão. Foi realizada uma filtragem para a exclusão dos transcritos sem valor de expressão (NA) e os valores foram transformados em  $\log_2$ . Obtiveram-se 6.876 genes para Caco-2/EH41 e 7.555 genes para Caco-2/Ec472, anotados no banco de dados *Gene Ontology* (GO). Após essa etapa, foi feita a checagem da variação dos valores de expressão através da plotagem, em gráfico boxplot, de cada amostra em cada um dos grupos Caco-2/EH41 e Caco-2/Ec472. Os dados de expressão gênica dos grupos Caco-2/EH41 e Caco-2/Ec472 mostraram que os valores de expressão variaram significativamente entre as amostras (Figura 7A e 8A), desta forma os dados foram normalizados para cada um dos grupos (Figura 7B e 8B). A matriz de dados normalizados e em  $\log_2$  foi utilizada para a composição das matrizes de dados de expressão das séries temporais para cada um dos grupos (Caco-2/EH41 e Caco-2/Ec472).

Todos os dados brutos dos microarranjos foram depositados no banco de dados públicos de genômica funcional *Gene Expression Omnibus* (GEO) (<http://www.ncbi.nlm.nih.gov/geo>), *National Center for Biotechnology Information* (NCBI), sob o número de acesso GSE104492.



**Figura 7** - Gráfico boxplot dos valores de expressão dos microarranjos de DNA das amostras T0 a T24 de Caco-2/EH41. **A** - valor de expressão em  $\log_2$ ; **B** - valor de expressão em  $\log_2$  normalizados



**Figura 8** - Gráfico boxplot dos valores de expressão dos microarranjos de DNA das amostras T0 a T24 de Caco-2/Ec472. **A** - valor de expressão em  $\log_2$ ; **B** - valor de expressão em  $\log_2$  normalizados

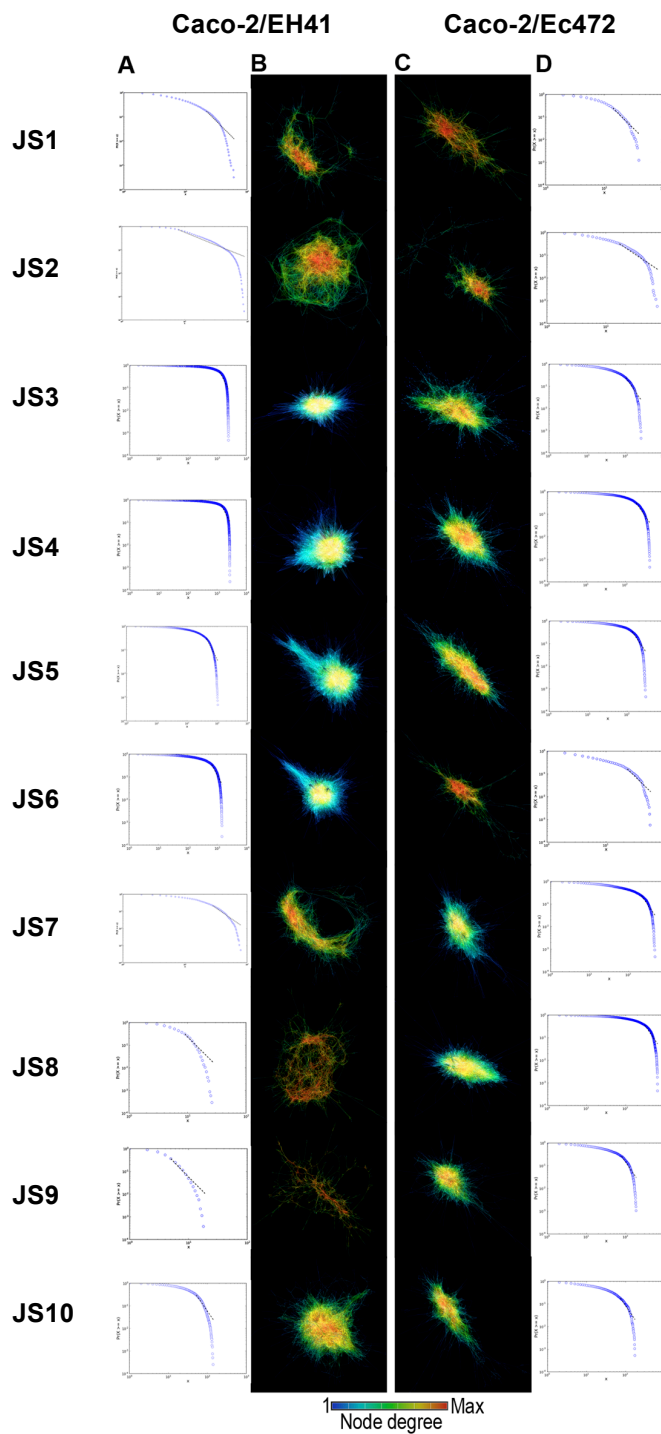
## 4.2 Análise da variação topológica de redes de coexpressão gênica

A análise da variação topológica de RCG para os grupos Caco-2/EH41 e Caco-2/Ec472 foi realizada através da comparação de 10 redes construídas a partir dos dados de expressão gênica diferencial e agrupadas em 10 janelas sobrepostas (ver Figura 5).

### 4.2.1 Redes de Caco-2/EH41

A análise comparativa das 10 janelas - JS1 a JS10 - revelou 4.285 genes diferencialmente expressos (DE) e anotados no banco GO. Esses genes foram utilizados para a construção das redes de coexpressão gênica (RCG) para cada uma das janelas, adotando um limiar de correlação gene-gene de 0,984 para todas as janelas. Esse limiar de correlação foi definido após ser verificado que a topologia da rede JS1 não se alterava dentro de um intervalo de correlação de Pearson (0,700 a 0,999).

A topologia das redes foi avaliada quanto ao tipo de distribuição de número de ligações por nó. Na Figura 9 é possível observar uma intensa variação topológica das redes ao longo do tempo. As janelas JS1 e JS2 apresentaram redes com status *scale-free*, como mostram os gráficos da distribuição de *node degree* ( $P(k)$ ) versus *node degree* ( $k$ ) e pela topologia das redes (Figura 9 A e B). Nas quatro janelas seguintes (JS3 a JS6) essas redes perdem esse status (muitos genes com muitas ligações). É interessante notar que nas próximas janelas JS7 a JS9 as redes voltaram a ter características *scale-free* e, finalmente, em JS10, a rede novamente não apresenta mais esse status.



**Figura 9** - Redes de coexpressão gênica dos grupos Caco-2/EH41 (**B**) ou Caco-2/Ec472 (**C**) obtidas das janelas JS1 a JS10. Teste Kolmogorov-Smirnov (K-S) para avaliar a topologia das redes (status *scale-free*) para as redes do grupo Caco-2/EH41 (**A**) e do grupo Caco-2/Ec472 (**D**)

#### 4.2.2 Redes de Caco-2/Ec472

A análise comparativa das 10 janelas JS1 a JS10 revelou 2.314 genes DE anotados no banco GO. Esses genes foram utilizados para a construção das redes para cada uma das janelas e adotou-se um limiar de correlação gene-gene de 0,990 para todas as janelas (após verificar-se que a topologia da rede JS1 não se alterava dentro de um intervalo de correlação de Pearson de 0,700 a 0,999).

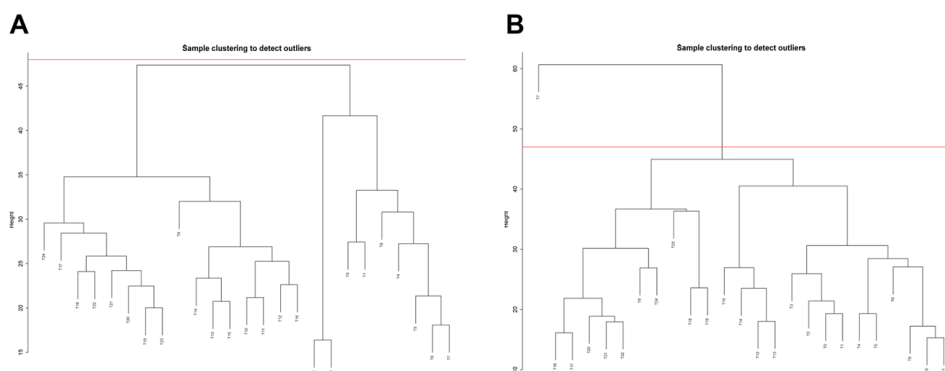
As Figuras 9 C e D mostram a variação topológica das redes JS1 a JS10 de Caco-2/Ec472. As janelas JS1, JS2 e JS6 apresentaram redes com características *scale-free*. Já nas demais redes (JS3 a JS5 e JS7 a JS10) essas redes não apresentam mais o status *scale-free*. Estes resultados mostram que as redes após uma hora de interação com Ec472/01 apresentaram pouca variação topológica.

#### 4.3 WGCNA

WGCNA foi utilizada para construção de RCG dos grupos Caco-2/EH41 e Caco-2/Ec472 para a identificação de módulos transcricionais fortemente correlacionados com três janelas justapostas durante o período de interação de Caco-2 com STEC (Figura 6).

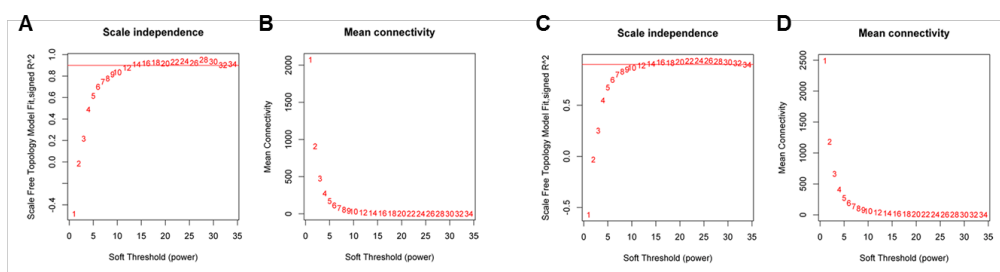
##### 4.3.1 Redes e identificação de módulos transcricionais

A primeira etapa foi a análise da matriz de dados de expressão gênica para identificar se haviam genes com valores discrepantes (*outliers*) tanto para o grupo Caco-2/EH41 quanto para o grupo Caco-2/Ec472. Após a filtragem obteve-se 6.876 genes para Caco-2/EH41 e 7.555 genes para Caco-2/Ec472 os quais foram utilizados para as análises subsequentes. Em seguida, foi feito o agrupamento das amostras, sendo que aqui uma amostra revelou-se como *outlier* (amostra T7 em Caco-2/Ec472) e foi descartada da análise (Figura 10).



**Figura 10** - Dendrograma das amostras para exclusão de *outliers*. Dendrograma do grupo Caco-2/EH41 (A) e do grupo Caco-2/Ec472 (B). A linha vermelha indica a linha de corte adotada para os grupos Caco-2/EH41 e Caco-2/Ec472

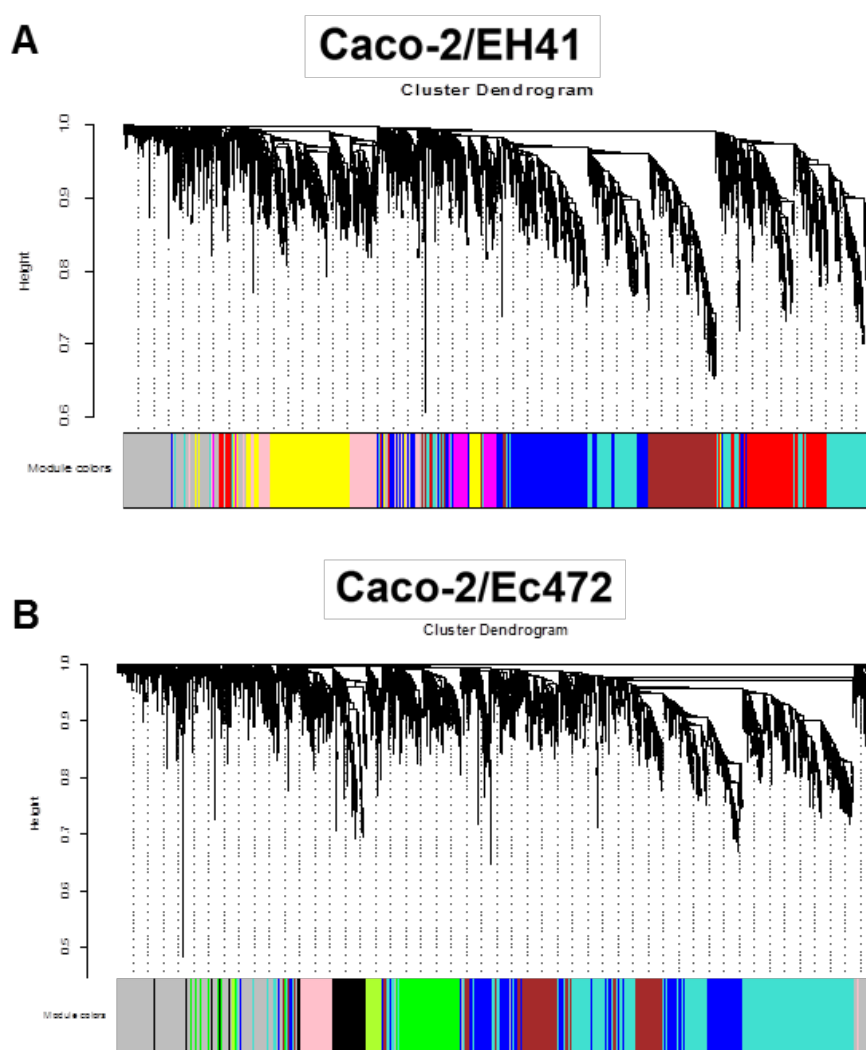
Após a exclusão da amostra *outlier*, duas redes foram construídas, uma para Caco-2/EH41 e outra para Caco-2/Ec472. A fim de obter uma rede *scale-free*, escolheu-se, para as duas redes, o valor  $\beta = 16$ , onde o parâmetro  $R^2$  apresentava valor acima de 0,90 (Figura 11).



**Figura 11** - Curva da variação topológica da rede para diferentes valores  $\beta$ . Curva de dispersão onde se mostra a distribuição dos valores do índice  $R^2$  (eixo y) para diferentes valores  $\beta$  (eixo x) obtidos para as redes dos grupos Caco-2/EH41 (A) e Caco-2/Ec472 (C). Gráficos mostrando a distribuição da média de conectividade dos nós da rede (eixo y) para diferentes valores  $\beta$  (eixo x) das redes dos grupos Caco-2/EH41 (B) e Caco-2/Ec472 (D). A linha vermelha indica a linha de corte ( $R^2=0,90$ )

Definido o parâmetro para construção do dendrograma dos genes, o próximo passo foi a escolha do tamanho mínimo do módulo e do limiar de corte dos ramos para a obtenção dos módulos transcricionais com o objetivo de obter aproximadamente 10 módulos, conforme recomendado pelo tutorial do WGCNA. Assim, adotou-se um limiar de corte de 0,25 e tamanho mínimo do

módulo de 50 genes para o grupo Caco-2/EH41 ou de 60 genes para o grupo Caco-2/Ec472. Foram identificados 11 módulos eigengenes para cada dendrograma, identificados por cores distintas (Figura 12). O tamanho dos módulos variou de 112 genes (módulo *green-yellow*) a 1.010 genes (módulo *turquoise*) para o grupo Caco-2/EH41 e de 118 (módulo *green-yellow*) a 2.472 genes (módulo *turquoise*) para o grupo Caco-2/Ec472 (Tabela 2).



**Figura 12** - Dendrograma dos genes e módulos transcricionais. **A** - Dendrograma dos grupos Caco-2/EH41 (**A**) e Caco-2/Ec472 (**B**). A barra de cores mostra a atribuição do módulo



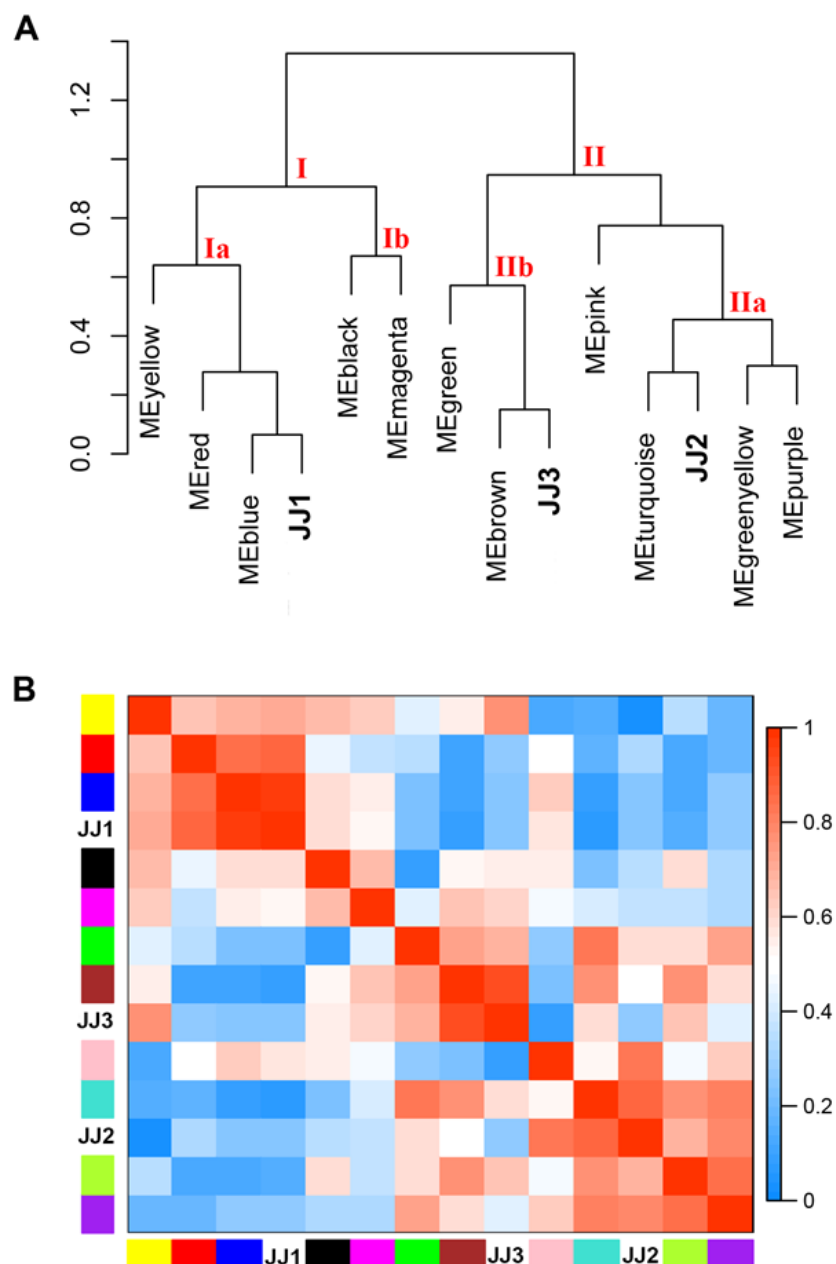
**Tabela 2 - Distribuição do número de genes nos 11 módulos das redes de Caco-2/EH41 e de Caco-2/Ec472**

Módulo	Caco-2/EH41	Caco-2/Ec472
	número de genes	
1 <i>turquoise</i>	1010	2472
2 <i>blue</i>	901	654
3 <i>brown</i>	769	507
4 <i>yellow</i>	516	485
5 <i>green</i>	442	440
6 <i>red</i>	368	404
7 <i>black</i>	241	252
8 <i>pink</i>	208	250
9 <i>magenta</i>	167	141
10 <i>purple</i>	121	133
11 <i>green-yellow</i>	112	118
- <i>grey</i>	2021	1699

#### 4.3.2 Caracterização funcional dos módulos transcricionais

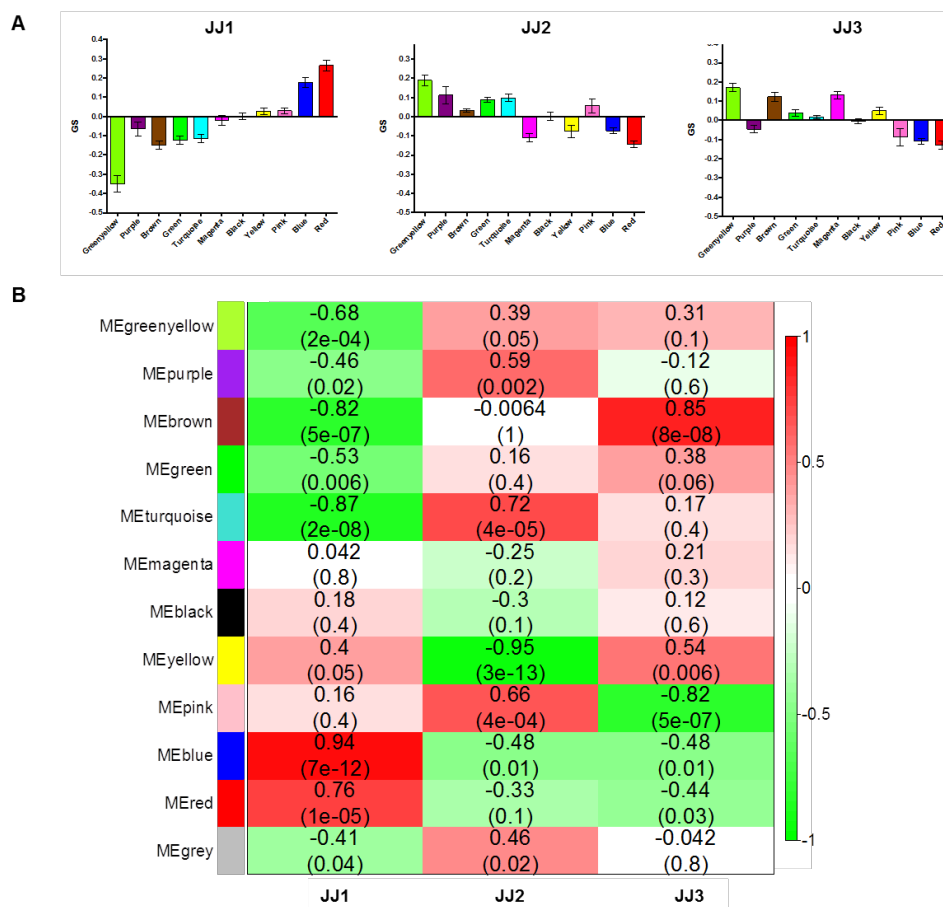
##### 4.3.2.1 Rede Caco-2/EH41

O dendrograma de módulos eigengene para Caco-2/EH41 apresentou dois grandes grupos ou meta-módulos, denominados I e II na Figura 13A. É interessante notar que o meta-módulo I correlaciona-se com JJ1 enquanto que o meta-módulo II correlaciona-se com as demais janelas (JJ2 e JJ3). Mais ainda, o meta-módulo II subdivide-se em dois grupos IIa e IIb. O meta-módulo IIa está relacionado com JJ2 enquanto que o IIb com JJ3. Este resultado indica que a rede apresenta dois grandes meta-módulos distintos durante as três horas de interação. O meta-módulo I está relacionado com a primeira hora de interação e o meta-modulo II com os intervalos após uma hora de interação enterócito-patógeno. A Figura 13B mostra o *heatmap* obtido para os módulos e janelas, onde a correlação forte está indicada em vermelho e a fraca em azul. Obteve-se alta correlação do módulo *blue* com JJ1, do módulo *turquoise* com JJ2 e módulo *brown* com JJ3. Estes resultados mostram que o perfil de expressão dos genes de cada um desses módulos está altamente correlacionado com as amostras de cada uma dessas janelas.



**Figura 13** - Rede eigengene dos módulos e das janelas do grupo Caco-2/EH41 e matriz de correlação entre os módulos e janelas. **A** - Dendrograma dos módulos eigengenes. **B** - Gráfico *heatmap* das correlações entre os módulos e as janelas. Cada linha e coluna no *heatmap* correspondem a um módulo eigengene (cada cor indica a cor do módulo) ou a janela (JJ1, JJ2 e JJ3). A barra de cor (à direita) indica o valor de correlação, onde a intensidade da cor varia conforme o valor de correlação

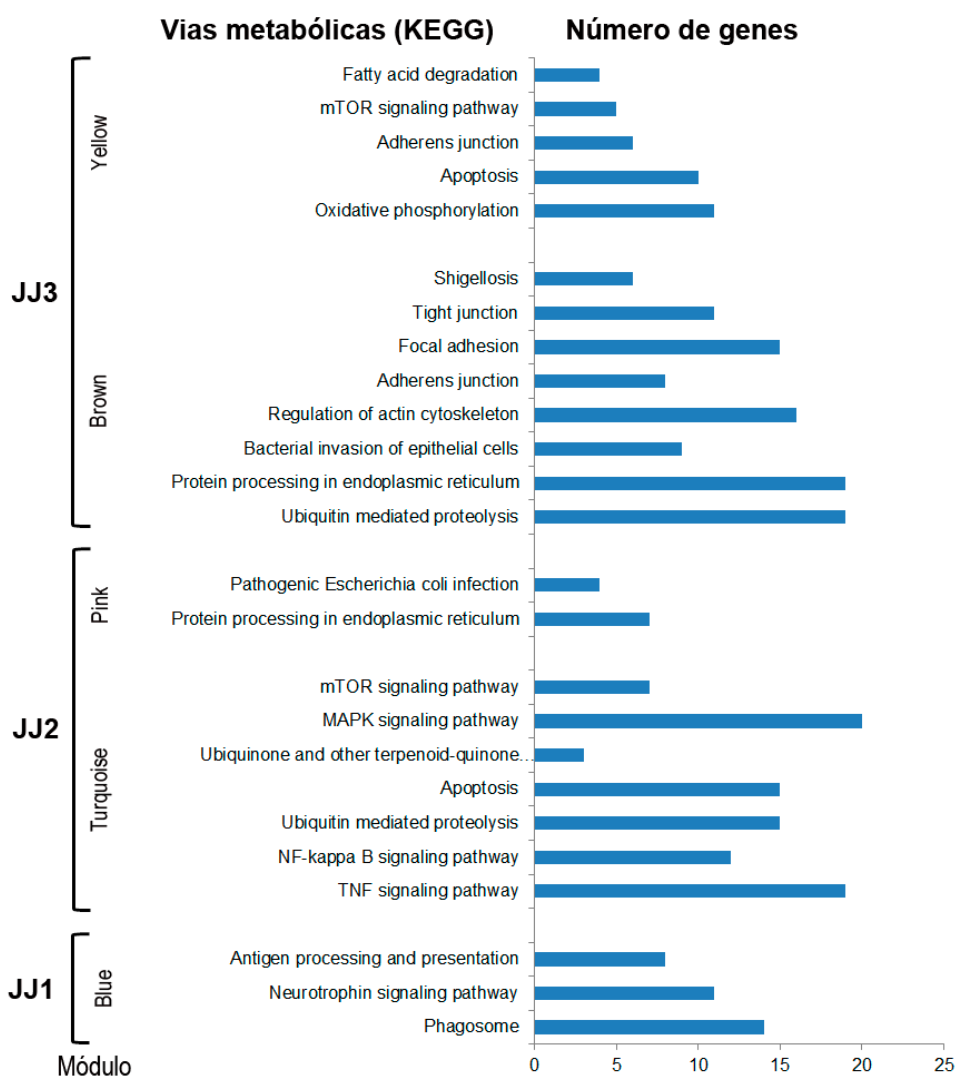
A partir dos resultados anteriormente descritos, a média dos valores GS (*Gene Significance*, i.e. a significância do perfil de expressão de um dado gene para uma determinada característica) dos genes de cada módulo foi plotada, bem como a análise de correlação do módulo com cada janela (Figura 14). A análise de distribuição dos valores GS dos genes de cada módulo mostrou que dois módulos possuem valores médios altos e positivos de GS: um deles é o módulo *red* para JJ1 e o outro é o módulo *green-yellow* para as janelas JJ2 e JJ3 (Figura 14A). Além disso, foi feita uma outra análise de correlação, com base na média dos valores absolutos de GS de cada módulo com cada janela. Esta análise revelou cinco módulos altamente associados e significativos para pelo menos uma janela ( $|r| > 0,80$ ;  $p < 0,01$ ) (Figura 14B). Os módulos *blue*, *brown* e *turquoise* estão associados com JJ1, mas somente o módulo *blue* está positivamente correlacionado. Somente um módulo, *yellow*, está negativamente associado a JJ2, enquanto que os módulos *brown* e *pink* estão respectivamente positiva e negativamente correlacionados com JJ3. Os módulos altamente correlacionados, porém negativamente associados às janelas, possuem associação positiva ( $r > 0,50$ ) e significativa com as janelas JJ2 (módulos *turquoise* e *pink*) e JJ3 (módulo *yellow*). Módulos com associação positiva possuem média dos valores GS dos genes positiva, sugerindo maior significância funcional para determinada característica fenotípica. Assim, adotou-se a correlação positiva módulo-janela para a análise funcional dos cinco módulos com as respectivas janelas.



**Figura 14** - Associação módulo-janela do grupo Caco-2/EH41. **A** - Distribuição das médias de GS dos módulos para cada janela (JJ1-3). **B** - Gráfico *heatmap* das correlações entre os módulos e as janelas. Cada linha corresponde a um módulo eigengene e cada coluna a uma janela. Cada célula contém o valor de correlação correspondente e o valor-p. A cor verde da legenda de cores (à direita) representa correlação negativa, enquanto a cor vermelha correlação positiva.

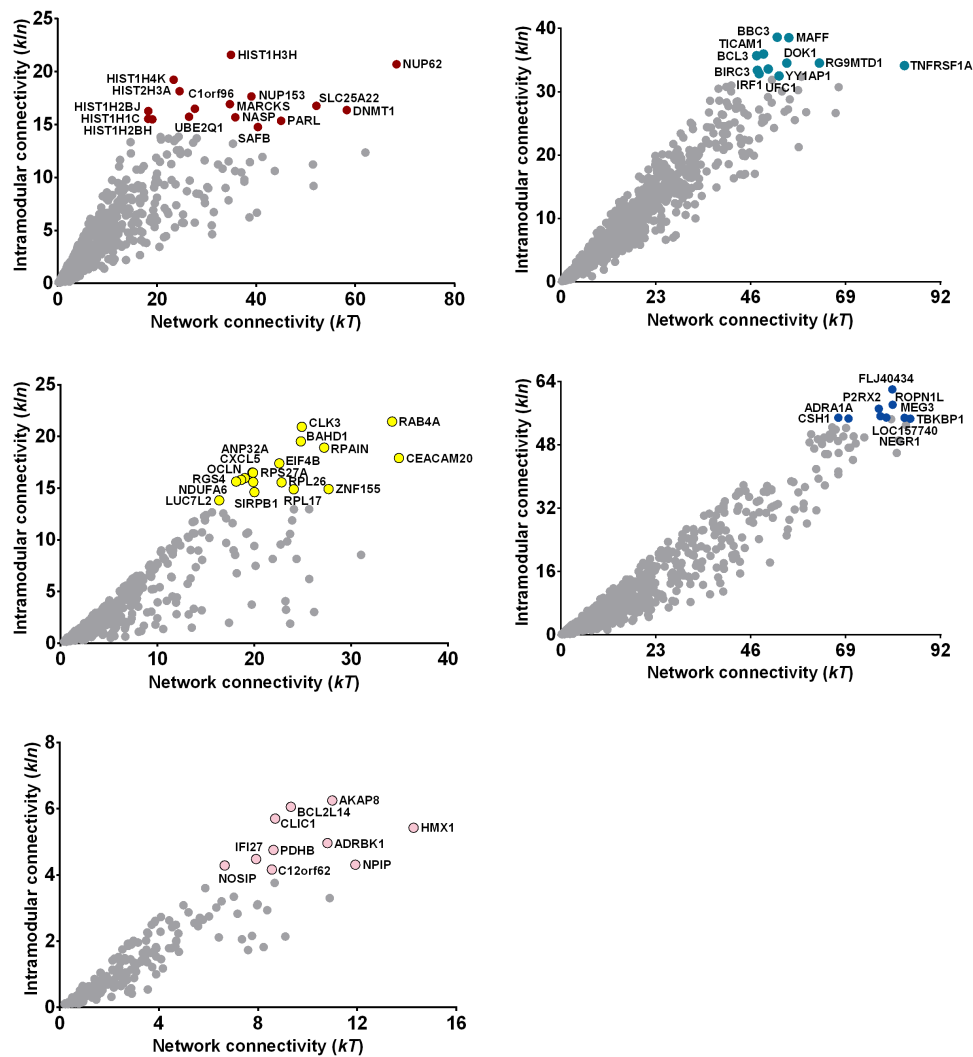
A primeira análise visou a uma caracterização global dos genes dos cinco módulos através da busca no banco KEGG de vias metabólicas utilizando uma ferramenta online Enrichr (<http://amp.pharm.mssm.edu/Enrichr/>). A Figura 15 mostra uma síntese das vias metabólicas possivelmente relacionadas com a resposta do enterócito a STEC. A lista completa das vias metabólicas encontradas ( $p < 0,05$ ) para cada módulo e os respectivos genes está apresentada no Anexo 3. A análise mostrou que muitos genes dos módulos

positivamente associados às janelas JJ2 e JJ3 estão envolvidos na resposta imune inata, apoptose e organização do citoesqueleto.

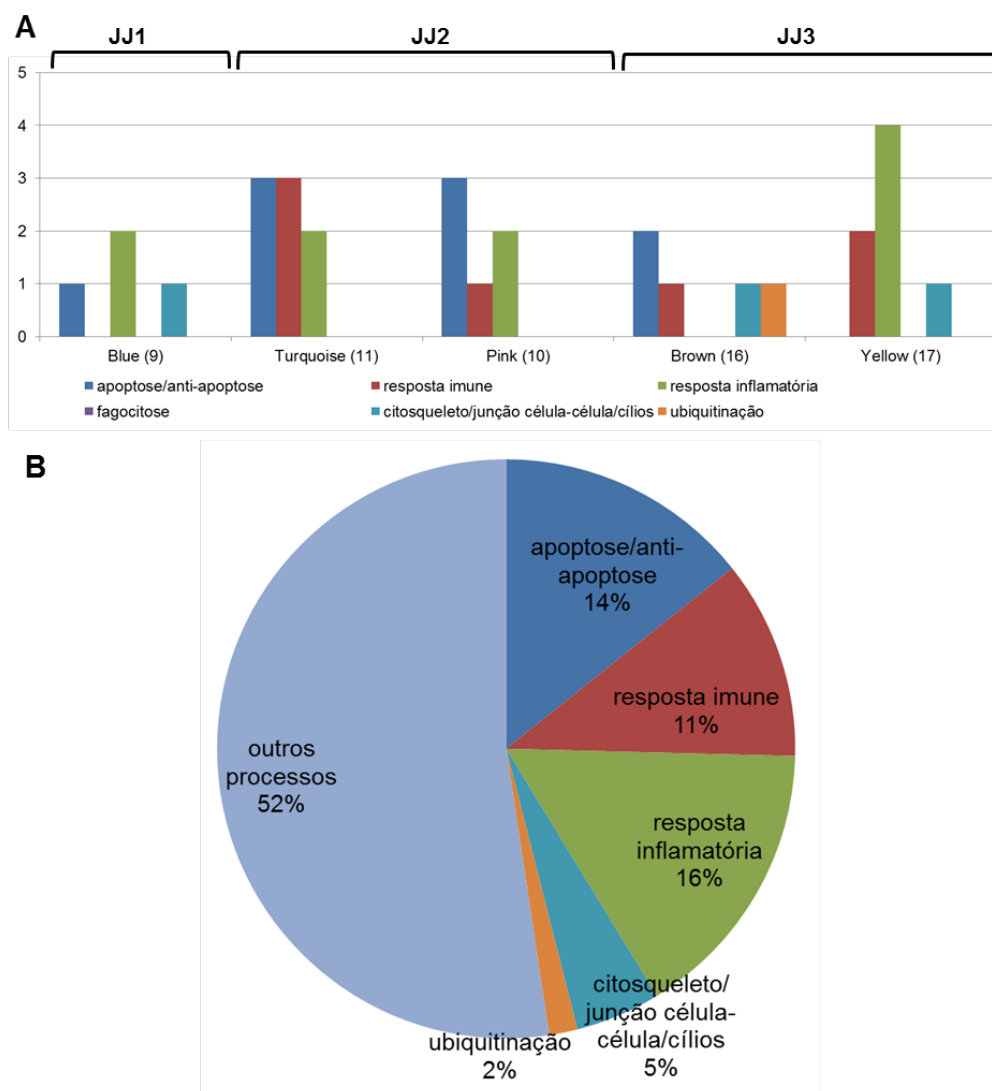


**Figura 15** - Histograma mostrando as principais vias metabólicas e número de genes dos módulos associados às janelas da rede do grupo Caco-2/EH41. Enriquecimento funcional dos genes baseado no banco KEGG (Enrichr,  $p < 0,05$ )

A próxima etapa consistiu em selecionar os hubs, i.e. os nós da rede que apresentaram alto  $k_{Total}$  e  $k_{Within}$  nesses módulos (Figura 16). Foi possível identificar um total de 63 hubs (Figura 17). Na Tabela 3 está listada a descrição das funções biológicas desses hubs. Na JJ1, dois hubs, *CSH1* e *TBKBP1*, codificam proteínas envolvidas na resposta inflamatória e um hub, *ADRA1A*, codifica uma proteína envolvida no processo apoptótico. Por outro lado, na JJ2 muitos hubs codificam proteínas envolvidas na resposta inflamatória (*TNFRSF1A* e *IRF1* hubs do módulo *turquoise*; *CLIC1* e *ADRBK1* do módulo *pink*) e imune (*TICAM1*, *BCL3* e *DOK1* do módulo *turquoise*; *AKAP8* do módulo *pink*) além de apoptose (*BBC3*, *UFC1* do módulo *turquoise*; *BCL2L14*, *IFI27* e *NPIP* do módulo *pink*) e anti-apoptose (*BIRC3*). Já na JJ3, três hubs (*BAHD1*, *CEACAM20* e *CXCL5*) do módulo *yellow* codificam proteínas envolvidas na resposta inflamatória e dois hubs, *NUP62* e *PARL*, do módulo *brown*, codificam proteínas envolvidas no processo apoptótico. Quatro hubs (*HIST1H2BJ* do módulo *brown*; *RAB4A*, *RPS27A* e *SIRPB1* do módulo *yellow*) codificam proteínas envolvidas na resposta imune e um hub, *MARCKS*, do módulo *brown*, codifica uma proteína envolvida no citoesqueleto (Figura 17, Tabela 3). A análise do perfil funcional de todos os hubs mostraram que quase metade desses genes podem estar envolvidos na resposta da Caco-2 à bactéria EH41. Vale a pena ressaltar que em JJ2, mais de 66% (14 de 21) dos hubs estão envolvidos na resposta imune e inflamação (Figura 19B). Estes resultados indicam que a STEC EH41 induz uma resposta imune, inflamatória e apoptótica da célula Caco-2 já na primeira hora de interação com as bactérias.



**Figura 16** - Gráfico de dispersão dos valores de  $k_{Total}$  ( $k_T$ ) vs  $k_{Within}$  ( $k_{In}$ ) dos genes dos módulos associados às janelas da rede do grupo Caco-2/EH41. As cores dos hubs correspondem às cores dos módulos. Os hubs estão identificados pelo símbolo dos genes



**Figura 17** - Análise funcional dos hubs dos módulos associados às janelas da rede do grupo Caco-2/EH41. **A** - Histograma das principais funções biológicas relacionadas com a resposta do enterócito após exposição à bactéria. Os números entre parênteses indicam o número total de hubs de cada módulo. **B** - Gráfico de setores do perfil funcional de todos os 63 hubs.



**Tabela 3 - Descrição da função biológica dos hubs da rede do grupo Caco-2/EH41**

Gene	Módulo	Gene Ontology		Descrição funcional
		Function	Process	
<b>FLJ40434</b>	Blue	ND	ND	desconhecido
<b>ROPN1L</b>		protein binding	epithelial cilium movement	movimento ciliar epitelial
<b>P2RX2</b>		extracellular ATP-gated cation channel activity	positive regulation of calcium-mediated signaling	sinalização
<b>LOC157740</b>	Blue	ND	ND	desconhecido
<b>NEGR1</b>		protein binding	locomotory behavior	ligação de proteína
<b>CSH1</b>		metal ion binding	JAK-STAT cascade involved in growth hormone signaling pathway	resposta inflamatória
<b>MEG3</b>	Blue	ND	ND	transcrição (Li et al., 2016)
<b>ADRA1A</b>		protein heterodimerization activity	apoptotic process; positive regulation of ERK1 and ERK2 cascade	apoptose
<b>TBKBP1</b>		protein binding	innate immune response	resposta inflamatória
<b>BBC3</b>	Turquoise	ATPase binding	intrinsic apoptotic signaling pathway	apoptose
<b>MAFF</b>		transcription factor activity	transcription from RNA polymerase II promoter	transcrição
<b>TICAM1</b>	Blue	signal transducer activity	positive regulation of I-kappaB kinase/NF-kappaB signaling; inflammatory response; positive regulation of autophagy	resposta imune inata
<b>BCL3</b>		transcription factor activity	I-kappaB kinase/NF-kappaB signaling; defense response to bacterium	resposta imune inata
<b>DOK1</b>		signal transducer activity	Ras protein signal transduction	resposta imune inata

continua

**Tabela 3 - Descrição da função biológica dos hubs da rede do grupo Caco-2/EH41**

			Transcrição	continuação
<b>RG9MTD1</b>	<i>RNA binding</i>	<i>positive regulation of mitochondrial translation</i>		
<b>TNFRSF1A</b>	<i>tumor necrosis factor binding</i>	<i>inflammatory response; positive regulation of I-kappaB kinase/NF-kappaB signaling</i>	resposta inflamatória	
<b>UFC1</b>	<i>protein binding</i>	<i>response to endoplasmic reticulum stress</i>	apoptose	
<b>BIRC3</b>	<i>ubiquitin-protein transferase activity</i>	<i>positive regulation of I-kappaB kinase/NF-kappaB signaling; regulation of innate immune response</i>	anti-apoptose	
<b>IRF1</b>	<i>RNA polymerase II core promoter proximal region sequence-specific DNA binding</i>	<i>positive regulation of interferon-beta production; regulation of innate immune response</i>	resposta inflamatória	
<b>YY1AP1</b>	<i>protein binding</i>	<i>regulation of cell cycle</i>	ciclo celular	
<b>AKAP8</b>	<i>Pink</i>	<i>cellular response to lipopolysaccharide; innate immune response</i>	resposta imune inata	
<b>BCL2L14</b>	<i>protein kinase binding</i>	<i>regulation of apoptotic process</i>	apoptose	
<b>CLIC1</b>	<i>NOT glutathione transferase activity</i>	<i>NOT glutathione metabolic process</i>	resposta inflamatória	
<b>HMX1</b>	<i>transcriptional repressor activity</i>	<i>negative regulation of transcription from RNA polymerase II promoter</i>	transcrição	
<b>ADRBK1</b>	<i>G-protein coupled receptor kinase activity</i>	<i>desensitization of G-protein coupled receptor protein signaling pathway</i>	resposta inflamatória	
<b>PDHB</b>	<i>contributes_to pyruvate dehydrogenase (NAD+) activity</i>	<i>mitochondrial acetyl-CoA biosynthetic process from pyruvate</i>	processo metabólico	

continua

**Tabela3 - Descrição da função biológica dos hubs da rede do grupo Caco-2/EH41**

continuação

<b>IFI27</b>	<i>RNA polymerase II activating transcription factor binding</i>	<i>activation of cysteine-type endopeptidase activity involved in apoptotic process</i>	apoptose
<b>NPIP</b>	ND	ND	apoptose (Garcia et al., 2009)
<b>NOSIP</b>	ND	<i>poly(A) RNA binding</i>	processo metabólico
<b>C12orf62</b>	ND	<i>COX14, cytochrome c oxidase assembly factor</i>	processo metabólico
<b>HIST1H3H</b>	Brown	<i>chromatin organization</i>	transcrição
<b>NUP62</b>	<i>histone binding</i> <i>ubiquitin binding</i>	<i>negative regulation of apoptotic process and programmed cell death</i>	apoptose
<b>HIST1H4K</b>	<i>histone binding</i>	<i>telomere organization</i>	transcrição
<b>HIST2H3A</b>	<i>histone binding</i>	<i>chromatin organization</i>	transcrição
<b>NUP153</b>	ND	<i>protein sumoylation</i>	transcrição
<b>MARCKS</b>	<i>actin filament binding</i>	<i>actin filament bundle assembly</i>	citoesqueleto
<b>SLC25A22</b>	<i>high-affinity glutamate transmembrane transporter activity</i>	<i>L-glutamate transmembrane transport, mitochondrial transport</i>	processo metabólico
<b>C1orf96</b>	<i>microtubule binding</i>	<i>regulation of mitotic spindle assembly</i>	divisão celular
<b>DNMT1</b>	<i>DNA-methyltransferase activity</i>	<i>transcription</i>	transcrição
<b>HIST1H2BJ</b>	<i>DNA binding</i>	<i>innate immune response in mucosa</i>	resposta imune inata
<b>UBE2Q1</b>	<i>ATP binding</i>	<i>protein ubiquitination</i>	ubiquitinação
<b>NASP</b>	<i>histone binding</i>	<i>DNA replication-dependent nucleosome assembly</i>	divisão celular
<b>HIST1H1C</b>	<i>chromatin DNA binding</i>	<i>nucleosome assembly</i>	divisão celular

continua



**Tabela 3 - Descrição da função biológica dos hubs da rede do grupo Caco-2/EH41**

				continuação
<b>HIST1H2BH</b>			<i>chromatin organization</i>	organização da cromatina
<b>PARL</b>	<i>DNA binding</i>		<i>negative regulation of intrinsic apoptotic signaling pathway</i>	apoptose
<b>SAFB</b>	<i>chromatin binding</i>		<i>chromatin organization</i>	organização da cromatina
<b>RAB4A*</b>	<i>GTPase activity</i>	Yellow	<i>Rab protein signal transduction; antigen processing and presentation</i>	resposta imune
<b>CLK3</b>	<i>protein binding</i>		<i>protein autophosphorylation</i>	ligação de proteína
<b>BAHD1</b>	<i>transcription regulatory region sequence-specific DNA binding</i>		<i>negative regulation of transcription, DNA-templated</i>	resposta inflamatória
<b>RPAIN</b>	ND		ND	desconhecido
<b>CEACAM20</b>	ND		ND	resposta inflamatória (Kotanio et al., 2016)
<b>EIF4B*</b>	<i>ribosomal small subunit binding</i>		<i>eukaryotic translation initiation factor 4F complex assembly</i>	translação
<b>CXCL5</b>	<i>chemokine activity</i>		<i>cell chemotaxis; inflammatory response</i>	resposta inflamatória
<b>ANP32A</b>	<i>protein binding</i>		<i>intracellular signal transduction</i>	sinalização
<b>OCLN</b>	<i>protein domain specific binding</i>		<i>cell-cell junction organization</i>	junção célula-célula
<b>RGS4</b>	<i>GTPase activator activity</i>		<i>positive regulation of GTPase activity</i>	GTPase
<b>RPS27A</b>	<i>metal ion binding</i>		<i>I-kappaB kinase/NF-kappaB signaling; innate immune response</i>	resposta inflamatória
<b>RPL26</b>	<i>RNA binding</i>		<i>translation</i>	translação
<b>NDUFA6</b>	<i>NADH dehydrogenase (ubiquinone) activity</i>		<i>response to oxidative stress</i>	processo metabólico

continua

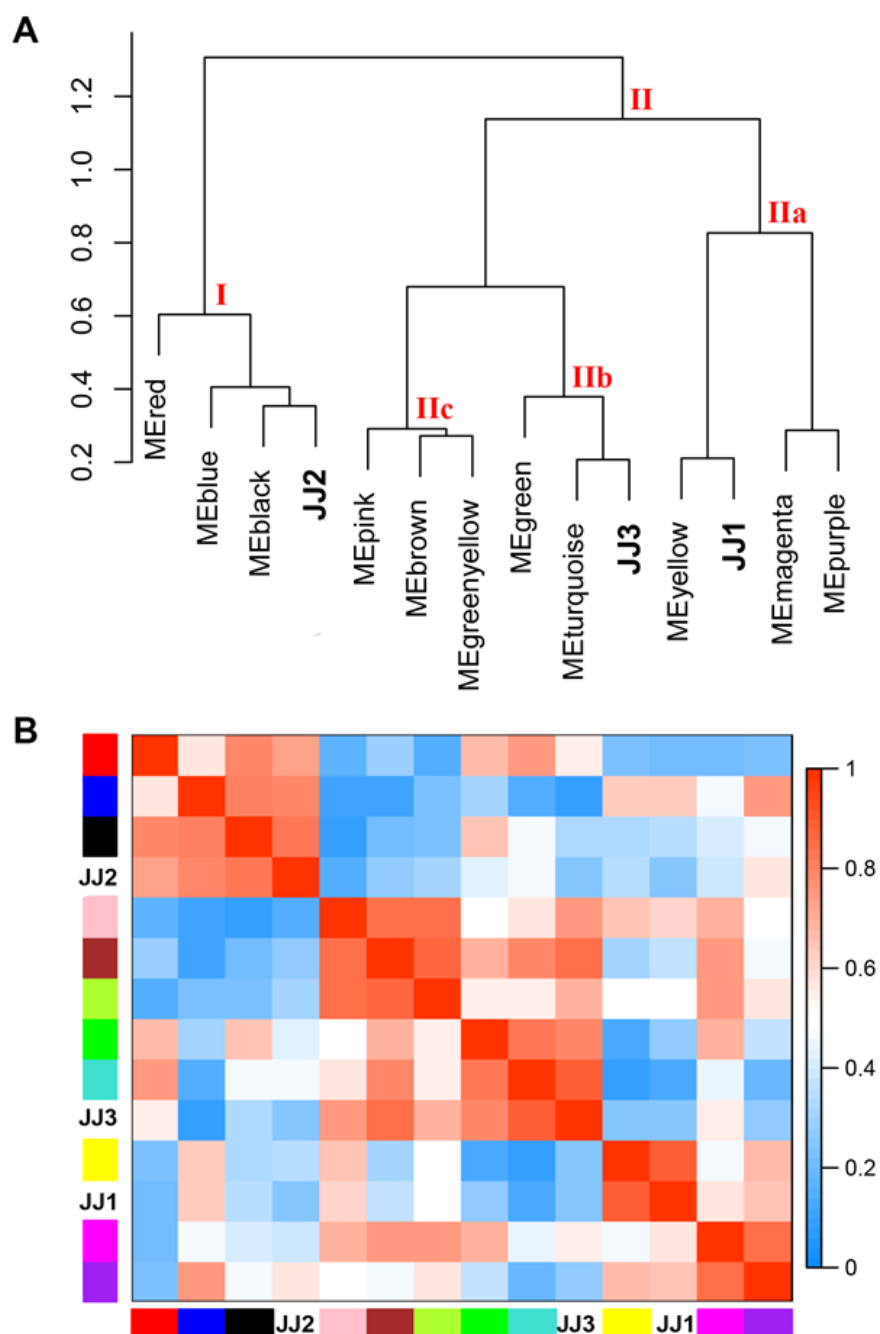
**Tabela 3 - Descrição da função biológica dos hubs da rede do grupo Caco-2/EH41**

			conclusão
<b>ZNF155</b>	<i>DNA binding</i>	<i>transcription</i>	transcrição
<b>RPL17</b>	<i>poly(A) RNA binding</i>	<i>translation</i>	translação
<b>SIRPB1</b>	<i>protein binding</i>	<i>signal transduction; innate immune response</i>	resposta imune inata
<b>LUC7L2</b>	<i>poly(A) RNA binding</i>	<i>mRNA splice site selection</i>	transcrição

ND, gene não descrito no GO; \*gene também é hub na rede Caco-2/Ec472.

#### 4.3.2.2 Rede Caco-2/Ec472

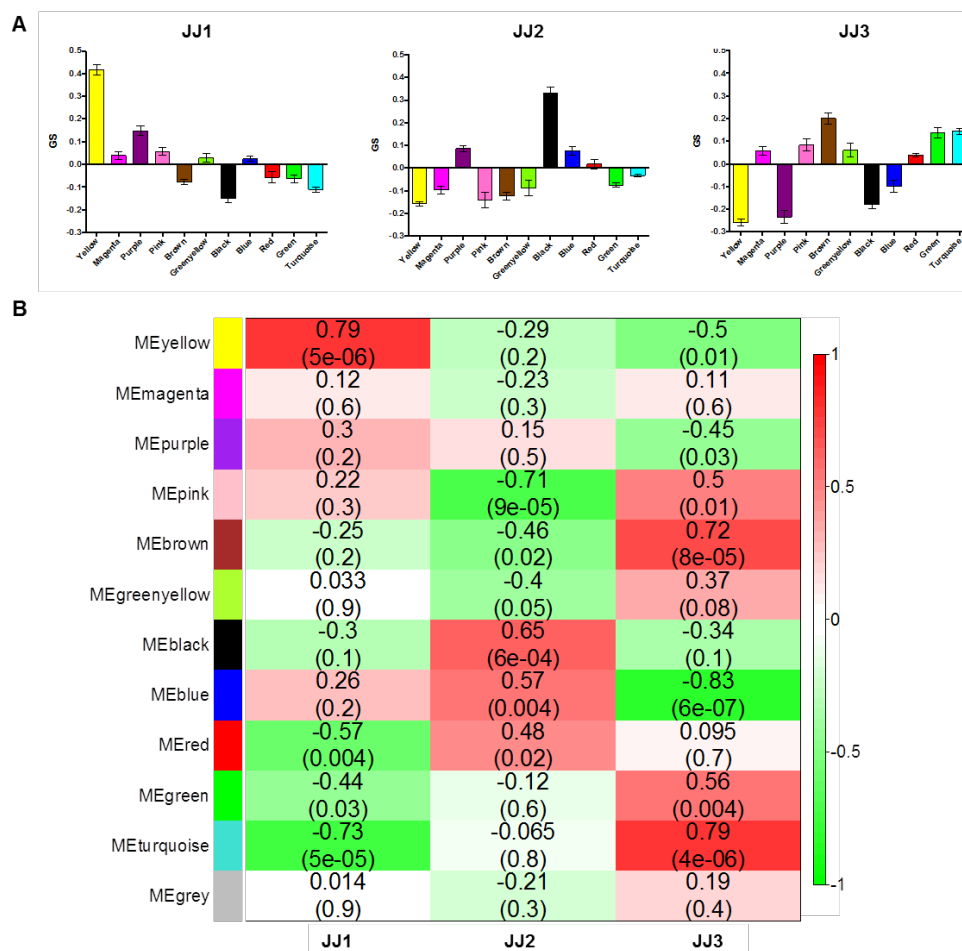
O dendrograma de módulos eigengene para Caco-2/Ec472 apresentou dois meta-módulos, I e II (Figura 18A). É interessante notar que o meta-módulo I possui somente três módulos e correlaciona com JJ2 enquanto que o meta-módulo II alberga a maioria dos módulos (8 de 11) e pode ser dividido em três meta-módulos IIa, IIb e IIc. O meta-módulo IIa está relacionado com JJ1 enquanto que o IIb com JJ3. Este resultado indica que a rede apresenta dois meta-módulos distintos durante as três horas de interação. O meta-módulo I está relacionado com a segunda hora de interação e o meta-módulo II com o início e após duas horas de interação enterócito-patógeno. A Figura 18B mostra o *heatmap* de correlação entre os módulos e as janelas. Obteve-se alta correlação do módulo *yellow* com JJ1, do módulo *black* com JJ2 e módulo *turquoise* com JJ3.



**Figura 18** - Rede eigengene dos módulos e das janelas de Caco-2/Ec472 e matriz de correlação entre os módulos e janelas. **A** - Dendrograma dos módulos eigengenes. **B** - Gráfico *heatmap* das correlações entre os módulos e as janelas. Cada linha e coluna no *heatmap* correspondem a um módulo eigengene (cada cor indica a cor do módulo) ou a janela (JJ1, JJ2 e JJ3). A barra de cor (à direita) indica o valor de correlação, onde a intensidade da cor varia conforme o valor de correlação

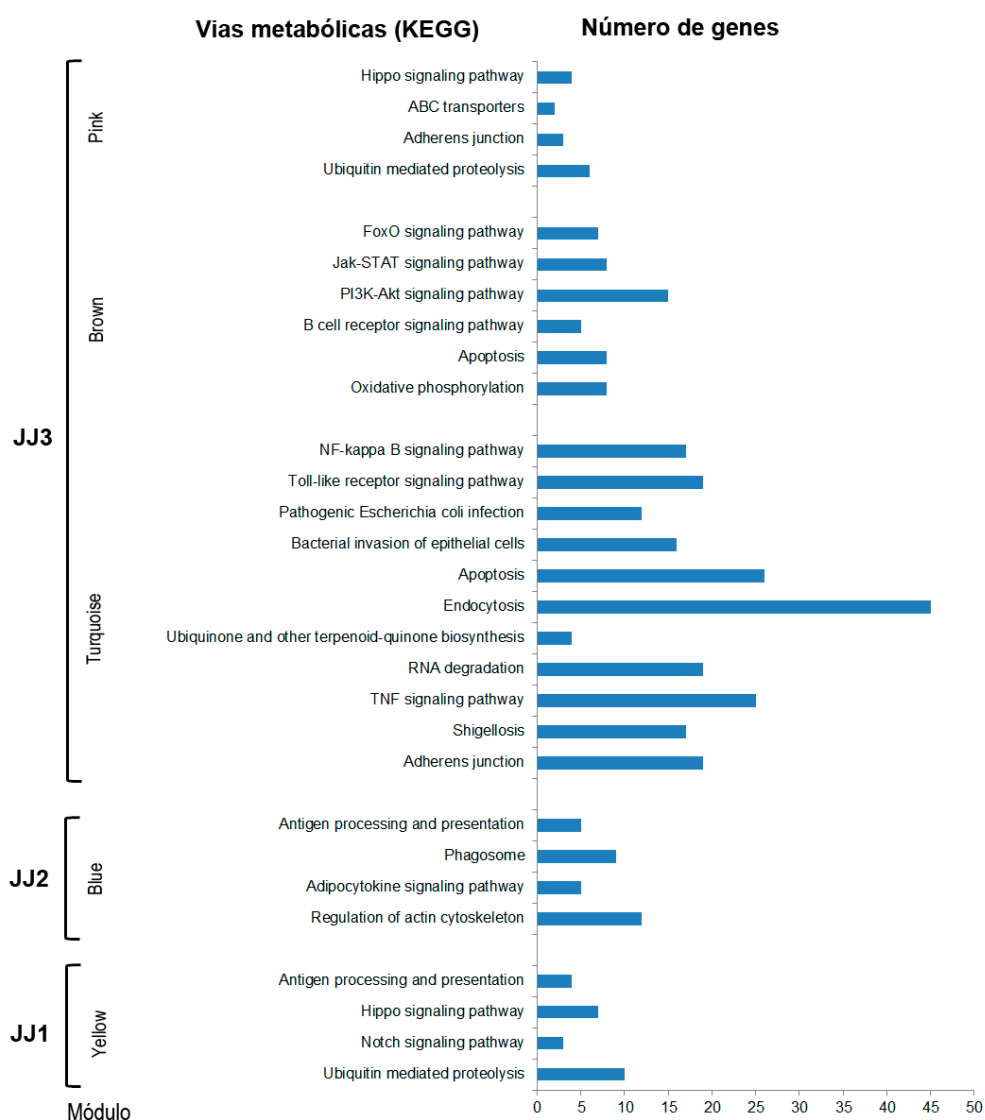
A Figura 19 mostra a média dos valores GS (*Gene Significance*) dos genes de cada módulo bem como a análise de correlação do módulo com cada janela. A análise de distribuição dos valores GS dos genes de cada módulo mostrou que pelo menos um módulo para cada janela possui alto valor médio positivo de GS, são eles: i) o módulo *yellow* para JJ1; ii) o módulo *black* para JJ2 e; iii) os módulos *brown*, *green* e *turquoise* para JJ3 (Figura 19A). Já a análise de correlação módulo-janela revelou cinco módulos altamente associados e significativos para pelo menos uma janela ( $|r| > 0,70$ ;  $p < 0,01$ ) (Figura 19B). Os módulos *yellow* e *turquoise* estão associados com JJ1, mas somente o módulo *yellow* está positivamente correlacionado. Somente um módulo, *pink*, está negativamente associado a JJ2, enquanto que três módulos estão associados com JJ3. Dois desses módulos - *brown* e *turquoise* - estão positivamente correlacionados e o terceiro módulo, *blue*, está negativamente correlacionado. Entretanto os módulos correlacionados negativamente possuem correlação positiva ( $r > 0,50$ ) e significativa com outra janela. Os módulos *blue* e *pink* estão positivamente correlacionados com as janelas JJ2 e JJ3 respectivamente. Desta forma, estes módulos foram selecionados para a caracterização funcional global dos genes - análise de enriquecimento baseado no banco KEGG - e a identificação de hubs através da categorização hierárquica dos genes obtido pela análise dos valores  $k_{Total}$  vs  $k_{Within}$  dos genes em um gráfico de dispersão.





**Figura 19** - Associação módulo-janela do grupo Caco-2/Ec472. **A** - Distribuição das médias de GS dos módulos para cada janela (JJ1-3). **B** - Gráfico *heatmap* das correlações entre os módulos e as janelas. Cada linha corresponde a um módulo eigengene e cada coluna a uma janela. Cada célula contém o valor de correlação correspondente e o valor-p. A cor verde da legenda de cores (à direita) representa correlação negativa, enquanto a cor vermelha correlação positiva.

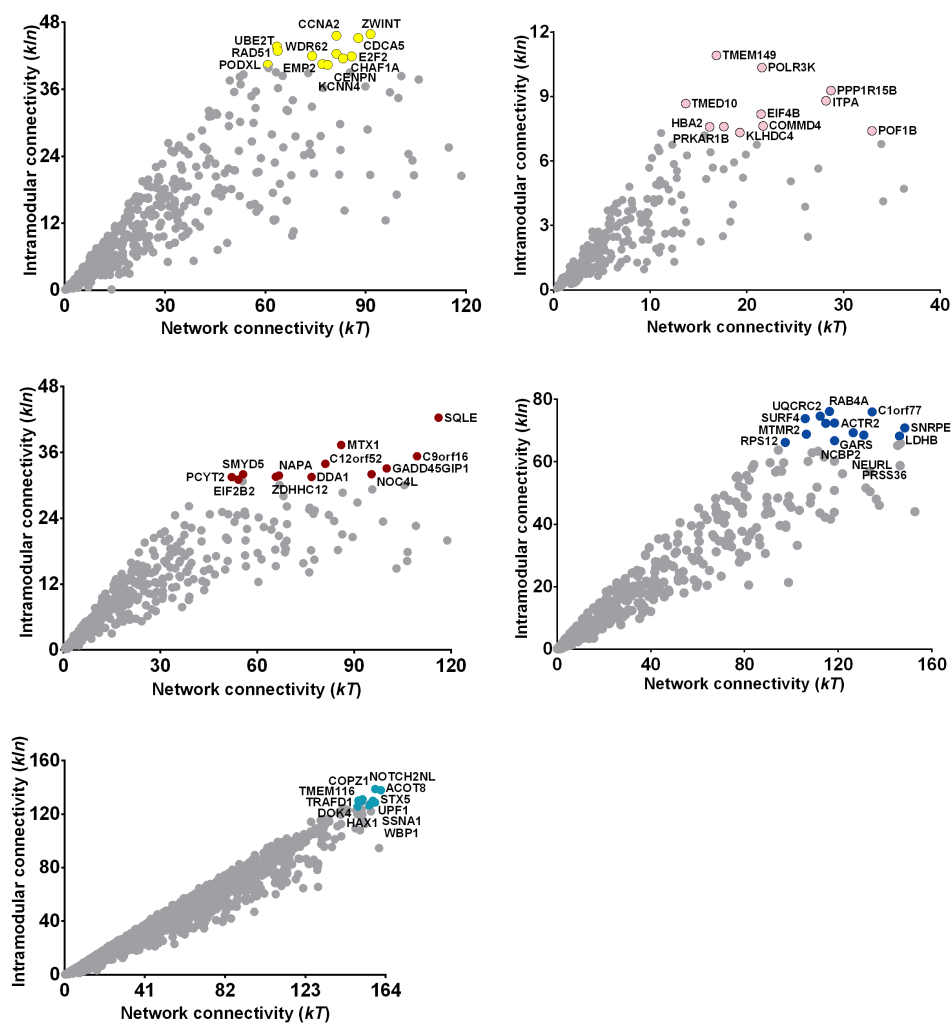
A Figura 20 mostra uma síntese das vias metabólicas possivelmente relacionadas com a resposta da Caco-2 à Ec472/01. A lista completa das vias metabólicas encontradas (banco KEGG,  $p < 0,05$ ) para cada módulo e os respectivos genes está no Anexo 4. A análise mostrou que muitos genes do módulo *turquoise* - positivamente associados a JJ3 - estão envolvidos na resposta imune.



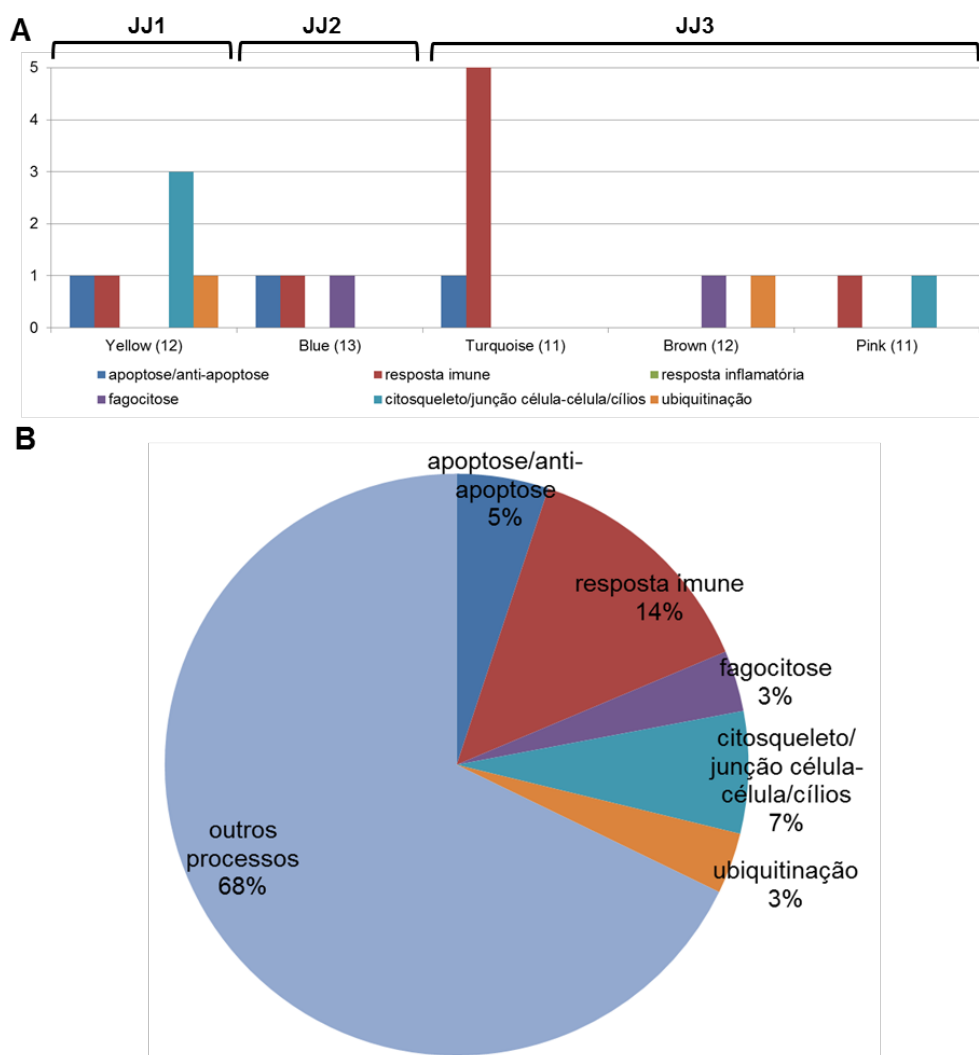
**Figura 20** - Histograma mostrando as principais vias metabólicas e o número de genes dos módulos associados às janelas da rede do grupo Caco-2/Ec472. Enriquecimento funcional dos genes baseado no banco KEGG (Enrichr,  $p < 0,05$ )

A Figura 21 mostra os gráficos de dispersão dos valores de  $k_{Total}$  vs  $k_{Within}$  dos genes de cada módulo. Desta forma, foi possível selecionar um total de 59 hubs. A Tabela 4 contém a descrição das funções biológicas desses hubs. Na JJ1, cinco hubs codificam proteínas que podem estar associados à

resposta do enterócito à Ec472/01. Três desses hubs - *CCNA2*, *EMP2* e *PODXL* - codificam proteínas envolvidas no citoesqueleto celular e outros dois hubs *KCNN4* e *E2F2* estão envolvidos respectivamente na resposta imune e apoptose. Na JJ2, um hub, *RAB4A*, que também é um hub da rede Caco-2/EH41 (janela JJ3) codifica uma proteína da família Ras de GTPase que está envolvida na resposta imune. Outro hub importante, *ACTR2*, codifica uma proteína estrutural do citoesqueleto e pode estar envolvido na fagocitose. Finalmente, o hub *NEURL* codifica uma proteína envolvida na apoptose. Já na JJ3, o processo mais importante associado parece ser a resposta imune, pois cinco hubs do módulo *turquoise* (*NOTCH2NL*, *ACOT8*, *UPF1*, *TRAFD1* e *DOK4*) e um hub, *POLR3K*, do módulo *pink* (Figura 22A, Tabela 4). A análise do perfil funcional de todos os hubs mostraram que apenas 32% (19 de 59) dos hubs podem estar envolvidos na resposta da Caco-2 à Ec472/01 (Figura 22B). Estes resultados mostraram que a célula Caco-2 tem uma resposta imune moderada quando desafiada por Ec472/01 e que esta resposta ocorre principalmente após duas horas de interação célula-bactéria. Esses resultados diferem dos obtidos para o grupo Caco-2/EH41, onde há muitos hubs envolvidos na resposta inflamatória já na primeira hora de interação.



**Figura 21** - Gráfico de dispersão dos valores de  $k_{Total}$  ( $k_T$ ) vs  $k_{Within}$  ( $k_{In}$ ) dos genes dos módulos associados às janelas da rede do grupo Caco-2/Ec472. As cores dos hubs correspondem às cores dos módulos. Os hubs estão identificados pelo símbolo dos genes



**Figura 22** - Análise funcional dos hubs dos módulos associados às janelas da rede do grupo Caco-2/Ec472. **A** - Histograma das principais funções biológicas relacionadas com a resposta do enterócito após exposição à bactéria. Os números entre parênteses indicam o número total de hubs de cada módulo. **B** - Gráfico de setores do perfil funcional de todos os 59 hubs.

**Tabela 4 - Descrição da função biológica dos hubs da rede do grupo Caco-2/Ec472**

Gene	Gene Ontology			Descrição funcional
	Módulo	Function	Process	
ZWINT	Yellow	protein binding	cell division	divisão celular
CCNA2		protein binding	Ras protein signal transduction	citoesqueleto (Loukil et al., 2015)
CDCA5		protein binding	mitotic cell cycle	divisão celular
UBE2T		ubiquitin protein ligase activity	DNA repair; protein ubiquitination	ubiquitinação
RAD51		protein binding	DNA repair; cellular response to DNA damage stimulus	divisão celular
CENPN		ND	mitotic cell cycle	ciclo celular
WDR62		protein binding	RNA splicing; centriole replication	divisão celular
E2F2		transcription factor binding	intrinsic apoptotic signaling pathway by p53 class mediator	apoptose
CHAF1A		unfolded protein binding	DNA repair; cell cycle	ciclo celular
EMP2		protein binding	actin filament organization; cell adhesion	citoesqueleto
KCNN4		calmodulin binding	defense response; immune system process	resposta imune
PODXL		protein binding	regulation of microvillus assembly	citoesqueleto
RAB4A*	Blue	GTPase activity	Rab protein signal transduction; antigen processing and presentation	resposta imune
C1orf77		protein binding	positive regulation of ATPase activity; positive regulation of cell proliferation	proliferação celular

continua

**Tabela 4 - Descrição da função biológica dos hubs da rede do grupo Caco-2/Ec472**

	<i>protein complex binding</i>	<i>protein processing involved in protein targeting to mitochondrion</i>	mitocôndria	continuação
<b>UQCRC2</b>		<i>retrograde vesicle-mediated transport, Golgi to ER</i>	Golgi/Reticulo endoplasmático; transporte de proteína	
<b>SURF4</b>	<i>protein binding</i>			
<b>ACTR2</b>	<i>structural constituent of cytoskeleton</i>	<i>Fc-gamma receptor signaling pathway involved in phagocytosis</i>	citoesqueleto; fagocitose (Bakovic et al., 2007)	
<b>NCBP2</b>	<i>protein binding</i>	<i>RNA metabolic process</i>	processo metabólico	
<b>SNRPE</b>	<i>protein binding</i>	<i>histone mRNA metabolic process</i>	transcrição	
<b>NEURL</b>	<i>ubiquitin-protein transferase activity; zinc ion binding</i>	<i>positive regulation of apoptotic process</i>	apoptose	
<b>MTMR2</b>	<i>protein binding</i>	<i>negative regulation of endocytosis</i>	transporte de endossomo	
<b>PRSS36</b>	<i>serine-type endopeptidase activity</i>	<i>proteolysis</i>	processo metabólico	
<b>LDHB</b>	<i>protein binding</i>	<i>oxidation-reduction process</i>	processo metabólico	
<b>GARS</b>	<i>ATP binding; protein dimerization activity</i>	<i>tRNA aminoacylation for protein translation</i>	translação	
<b>RPS12</b>	<i>structural constituent of ribosome</i>	<i>nuclear-transcribed mRNA catabolic process, nonsense-mediated decay</i>	transcrição	
<b>NOTCH2NL</b>	<i>Turquoise</i>	<i>Notch signaling pathway</i>	resposta imune	
<b>COPZ1</b>	<i>calcium ion binding; protein binding</i>	<i>ER to Golgi vesicle-mediated transport; toxin transport</i>	Golgi/Reticulo endoplasmático; transporte de proteína	
<b>TMEM116</b>	ND	ND	desconhecido	
<b>ACOT8</b>	<i>CoA hydrolase activity</i>	<i>acyl-CoA metabolic process; negative regulation of CD4 biosynthetic process</i>	resposta imune	

continua



**Tabela 4 - Descrição da função biológica dos hubs da rede do grupo Caco-2/Ec472**

				continuação
<b>STX5</b>	SNAP receptor activity; protein binding	ER to Golgi vesicle-mediated transport; regulation of Golgi organization	Golgi/Reticulo endoplasmático	
<b>HAX1</b>	interleukin-1 binding	cellular response to cytokine stimulus; negative regulation of apoptotic process	anti-apoptose	
<b>UPF1</b>	protein binding; zinc ion binding	DNA repair; cellular response to interleukin-1	resposta imune (Mao et al., 2017)	
<b>SSNA1</b>	identical protein binding	G2/M transition of mitotic cell cycle	ciclo celular	
<b>WBP1</b>	protein binding	ND	interação protein-protein	
<b>TRAFD1</b>	metal ion binding; protein binding	negative regulation of innate immune response; response to cytokine	resposta imune inata (Sanada et al., 2008)	
<b>DOK4</b>	protein binding	axon guidance	resposta imune	
<b>SQLE</b>	flavin adenine dinucleotide binding	oxidation-reduction process	processo metabólico	
<b>MTX1</b>	ND	protein targeting to mitochondrion	mitocôndria	
<b>C9orf16</b>	protein binding	ND	ligação de proteína	
<b>C12orf52</b>	protein binding	negative regulation of Notch signaling pathway	transcrição	
<b>GADD45GIP1</b>	protein binding	positive regulation of cell proliferation	proliferação celular	
<b>SMYD5</b>	metal ion binding	methylation	ciclo celular	
<b>NAPA</b>	soluble NSF attachment protein activity	ER to Golgi vesicle-mediated transport	Golgi/Reticulo endoplasmático	
<b>NOC4L</b>	protein binding	rRNA processing	transcrição	

continua



**Tabela 4 - Descrição da função biológica dos hubs da rede do grupo Caco-2/Ec472**

					continuação
<b>ZDHC12</b>	protein-cysteine S-palmitoyltransferase activity	protein palmitoylation	translação		
<b>DDA1</b>	protein binding	regulation of proteasomal ubiquitin-dependent protein catabolic process	ubiquitinação		
<b>PCYT2</b>	ethanolamine-phosphate cytidyltransferase activity	phosphatidylethanolamine biosynthetic process	fagocitose (Bakovic et al., 2007)		
<b>EIF2B2</b>	protein binding; contributes to translation initiation factor activity	cellular response to stimulus; positive regulation of GTPase activity	translação		
<b>TMEM149</b>	protein binding	ND	desconhecido		
<b>POLR3K</b>	zinc ion binding	innate 60mune response	resposta imune inata		
<b>PPP1R15B</b>	protein binding	ER overload response	Reticulo endoplasmático		
<b>ITPA</b>	metal ion binding	chromosome organization	organização da cromatina		
<b>TMED10</b>	protein binding	ER to Golgi vesicle-mediated transport; intracellular protein transport	transporte de proteína		
<b>EIF4B*</b>	protein binding	eukaryotic translation initiation factor 4F complex assembly	translação		
<b>COMMD4</b>	protein binding	regulation of transcription, DNA-templated	transcrição		
<b>PRKAR1B</b>	cAMP-dependent protein kinase regulator activity	activation of protein kinase A activity	processo metabólico		

continua

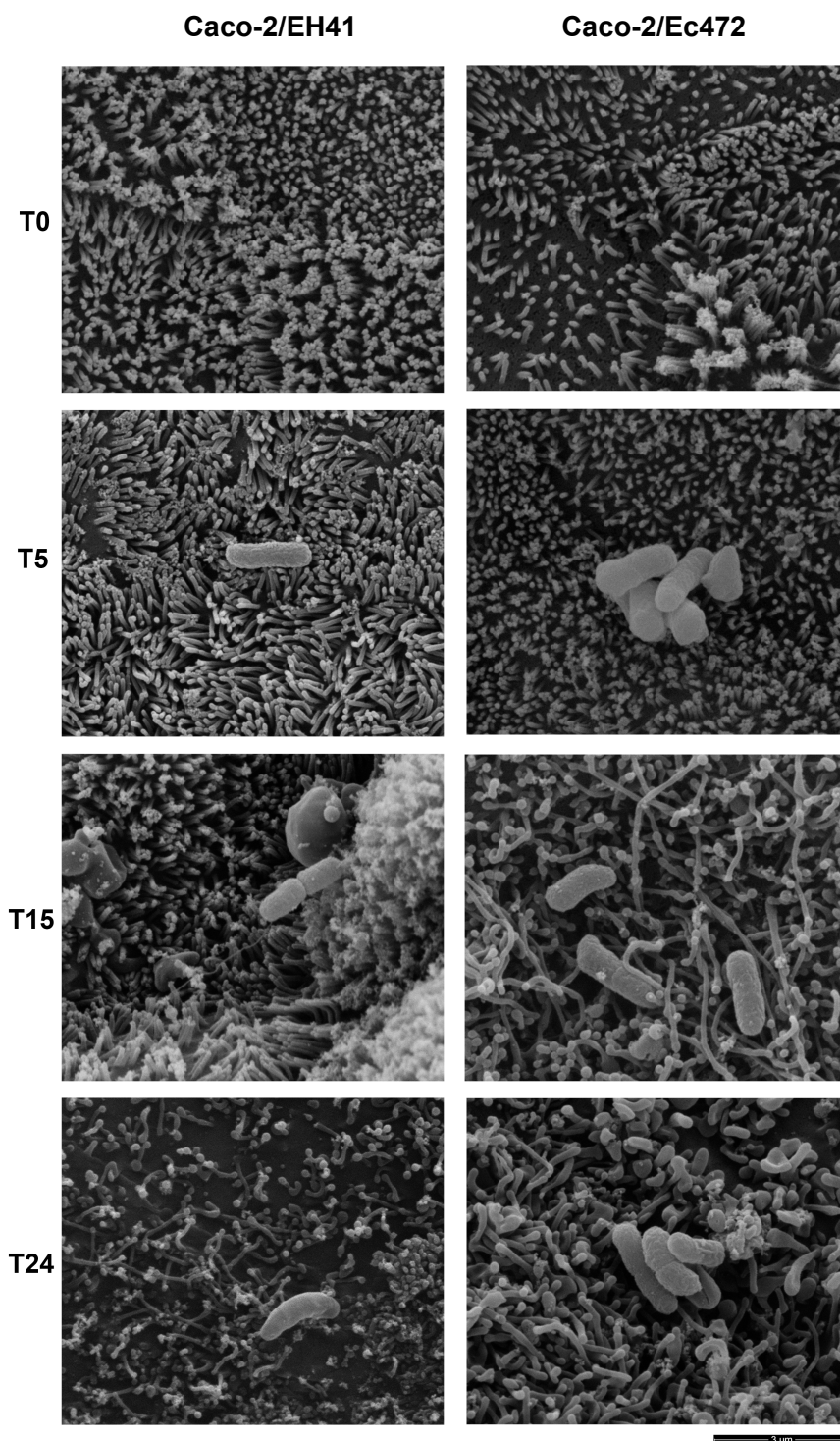
**Tabela 4 - Descrição da função biológica dos hubs da rede do grupo Caco-2/Ec472**

				conclusão
<b>HBA2</b>	<i>oxygen transporter activity</i>	<i>cellular oxidant detoxification</i>	processo metabólico	
<b>POF1B</b>	<i>actin filament binding</i>	<i>actin cytoskeleton organization</i>	citoesqueleto	
<b>KLHDC4</b>	<i>protein binding</i>	ND	ligação de proteína	

ND, gene não descrito no GO; \*gene também é hub na rede Caco-2/EH41.

#### 4.4 MEV

A Figura 23 mostra as imagens de MEV obtidas para Caco-2 durante a interação das células com as bactérias - isolados EH41 ou Ec472/01 - em diferentes intervalos de tempo (T0, T5, T15 e T24). A partir do intervalo T5 observam-se bactérias aderidas às células. Nos enterócitos expostos à EH41 observam-se alterações da superfície celular, como perda progressiva de microvilosidades a partir de T5, além de alterações na forma dos microvilos em T15 e, finalmente, uma perda extensa da microvilosidade em T24. As células Caco-2 expostas à Ec472 também apresentaram alterações morfológicas, entretanto distintas na forma e no tempo. Em T5 ainda não ocorre perda de microvilosidades, porém as células apresentam alterações morfológicas, como perda da orientação espacial. Em T15 há diminuição de microvilosidades e os microvilos tornam-se alongados; em T24, em linhas gerais, o quadro anteriormente descrito permanece.

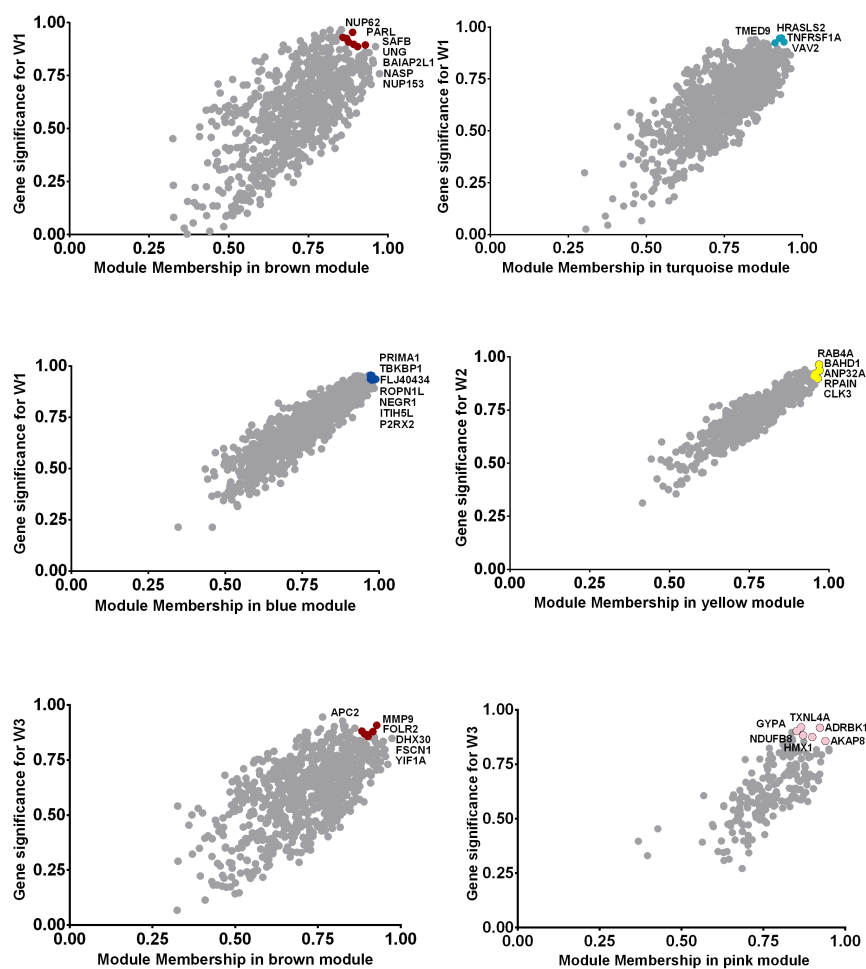


**Figura 23** - Imagens das células Caco-2 em diferentes intervalos de tempo (T0, T5, T15 e T24) durante a interação com EH41 ou Ec472/01 obtidas por MEV. A barra no canto inferior direito indica a escala de tamanho (3 μm)

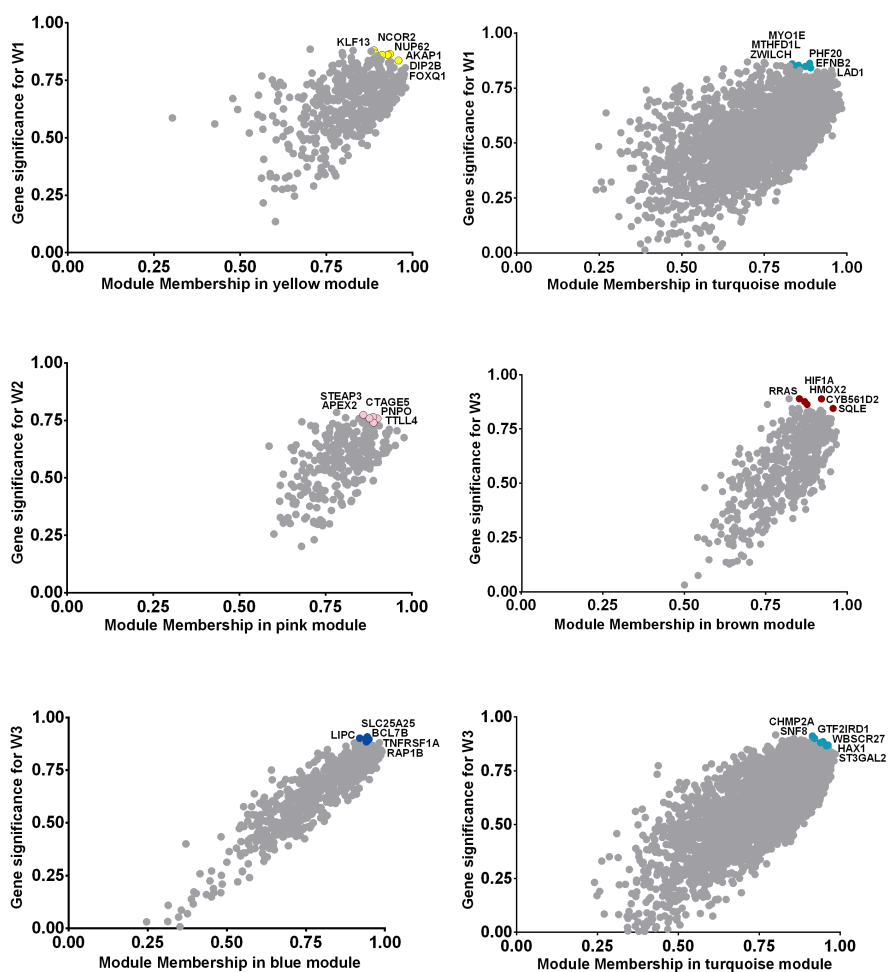
#### 4.5 Validação técnica dos microarranjos de DNA

Primeiramente foi necessário realizar uma análise estatística para a escolha dos genes a serem incluídos na validação. Módulos com alta correlação entre o gene e a janela foram selecionados para identificar os genes apresentando maior GS para pelo menos uma janela (Figuras 24 e 25). Foram selecionados 34 genes para o grupo Caco-2/EH41 e 29 genes para o grupo Caco-2/Ec472.

Foi realizada uma análise de variância das medianas dos dados de expressão gênica das amostras e comparado entre as janelas, utilizou-se o teste *Kruskal-Wallis* no programa *GraphPad Prism 6.0* (GraphPad, San Diego, CA) . Foram consideradas significativas as comparações que apresentaram  $p < 0.05$ . Assim foi possível identificar três genes com maior variância entre janelas para cada um dos grupos, Caco-2/EH41 ou Caco-2/Ec472. Os genes *TBKBP1*, *PRIMA1* e *NEGR1* para Caco-2/EH41, e os genes *TNFRSF1A*, *SLC25A25* e *BCL7B* para Caco-2/Ec472, foram utilizados para a validação técnica dos microarranjos de DNA.



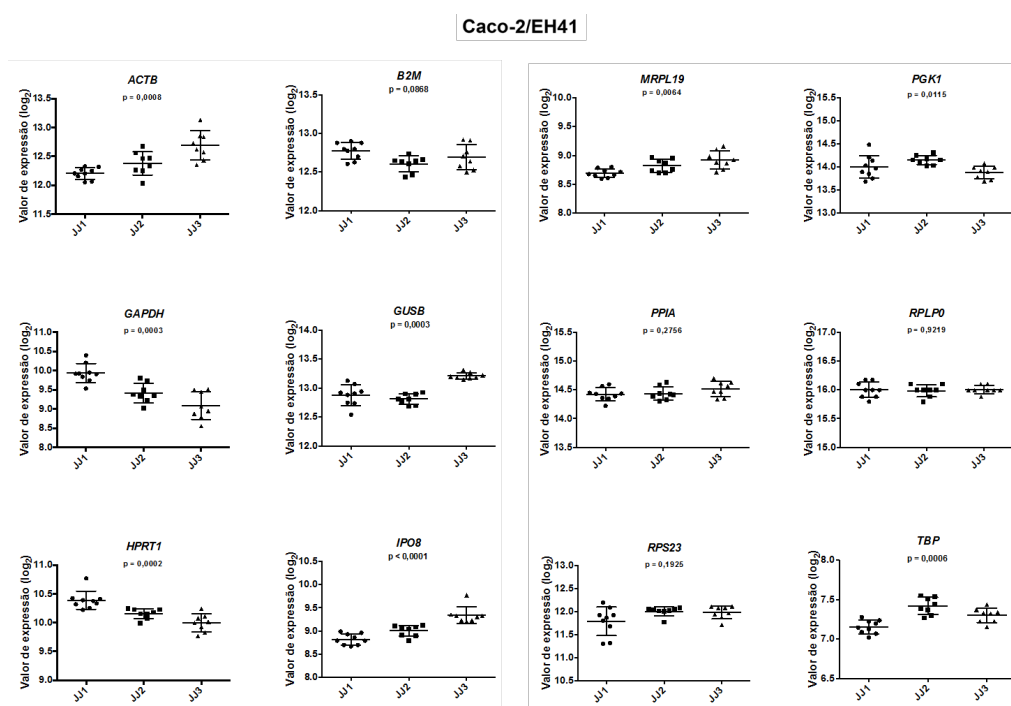
**Figura 24** - Gráfico de dispersão dos genes em cada módulo da rede do grupo Caco-2/EH41. Valores de *module membership* (eixo x) vs GS (eixo y) dos genes dos módulos altamente associados às janelas. As cores dos genes correspondem às cores dos módulos. Os genes estão identificados por seus símbolos



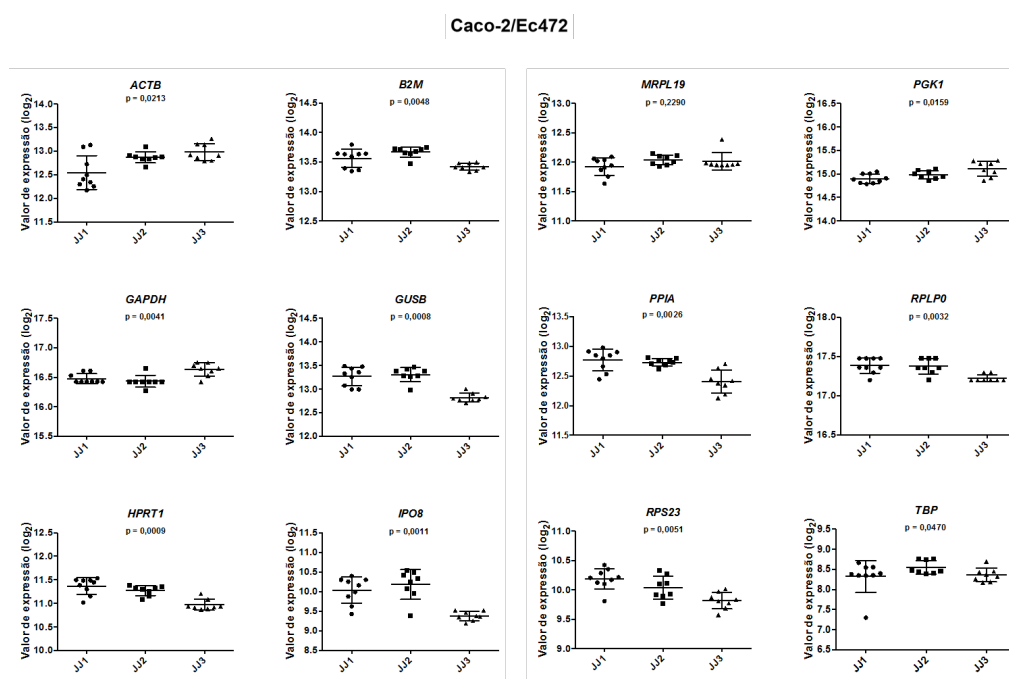
**Figura 25** - Gráfico de dispersão dos genes em cada módulo da rede do grupo Caco-2/Ec472. Valores de *module membership* (eixo x) vs GS (eixo y) dos genes dos módulos altamente associados às janelas. As cores dos genes correspondem às cores dos módulos. Os genes estão identificados por seus símbolos

O segundo passo foi selecionar um controle endógeno que, em cada grupo, não apresentasse variação de expressão ao longo do tempo. Foram inicialmente escolhidos 13 genes endógenos (*ACTB*, *B2M*, *GAPDH*, *GUSB*, *HPRT1*, *IPO8*, *MRPL19*, *PGK1*, *PPIA*, *RPLP0*, *RPS23* e *TBP*) já descritos para Caco-2 (Krzystek-Korpacka et al., 2016). A expressão desses genes, obtida por microarranjos de DNA de todas as amostras, foi plotada para se identificar genes utilizáveis como gene endógeno de referência. A análise de variância das medianas dos valores de expressão entre as janelas foi feita pelo teste de *Kruskal-Wallis*, considerando-se significativas as comparações que apresentaram  $p < 0,05$  (Figuras 26 e 27). Para o grupo Caco-2/EH41 verificou-se que os genes *PPIA*, *RPLP0* e *RPS23* não apresentam variação significativa entre as janelas. Para o grupo Caco-2/Ec472 foram os genes *MRPL19* e *TBP*. A expressão desses cinco genes foi então quantificada através da reação de qPCR. As Figuras 28 A e B mostram que os genes *RPLP0* e *PPIA* para o grupo Caco-2/EH41 e *MRPL19* e *TBP* para Caco-2/Ec472 não apresentaram variação de expressão entre as amostras testadas. Assim os genes *RPLP0* para o grupo Caco-2/EH41 e *MRPL19* para o grupo Caco-2/Ec472 foram utilizados como referência para a normalização das reações de qPCR para a validação técnica.

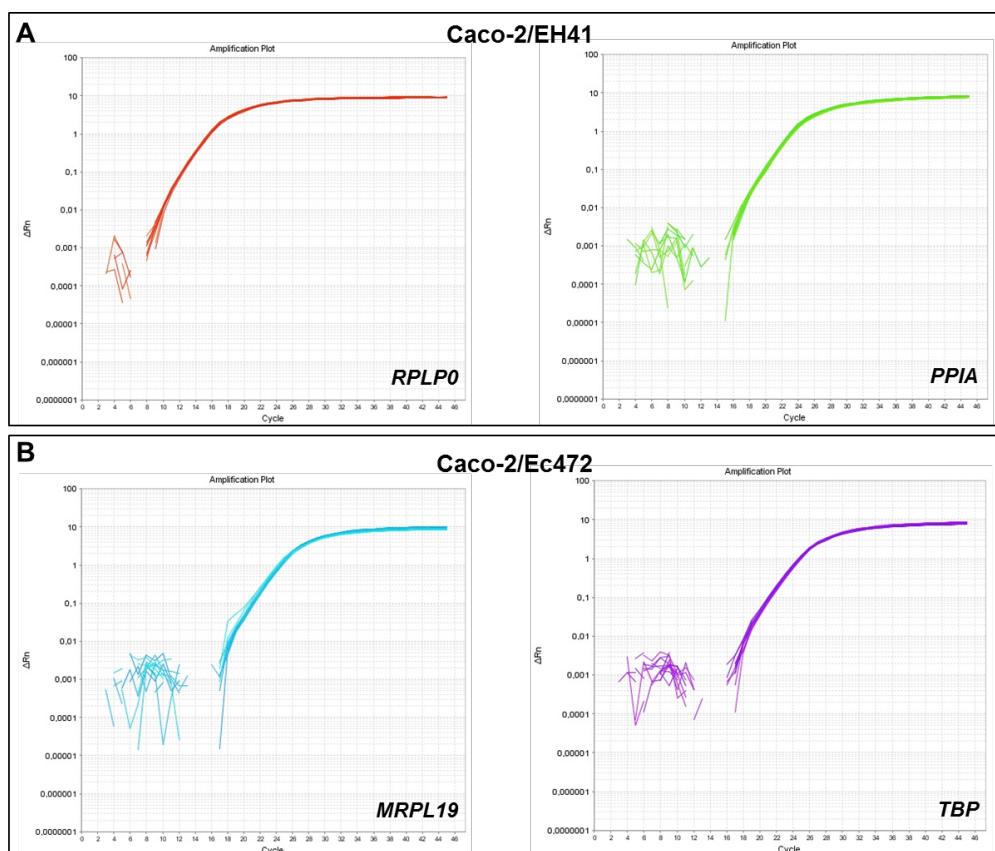




**Figura 26** - Mediana dos valores de expressão obtidos por microarranjos de DNA dos genes endógenos candidatos para a escolha de genes de referência para o grupo Caco-2/EH41. Teste de *Kruskal-Wallis*,  $p < 0,05$

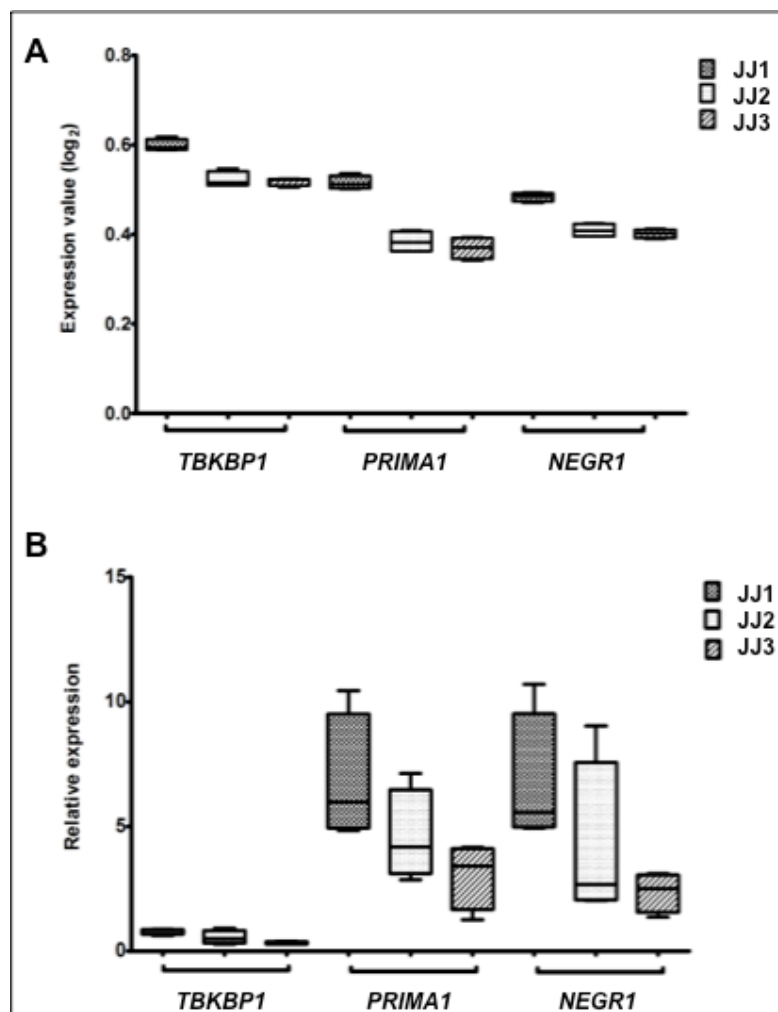


**Figura 27** - Mediana dos valores de expressão obtidos por microarranjos de DNA dos genes endógenos candidatos para a escolha de genes de referência para o grupo Caco-2/Ec472. Teste de *Kruskal-Wallis*,  $p < 0,05$

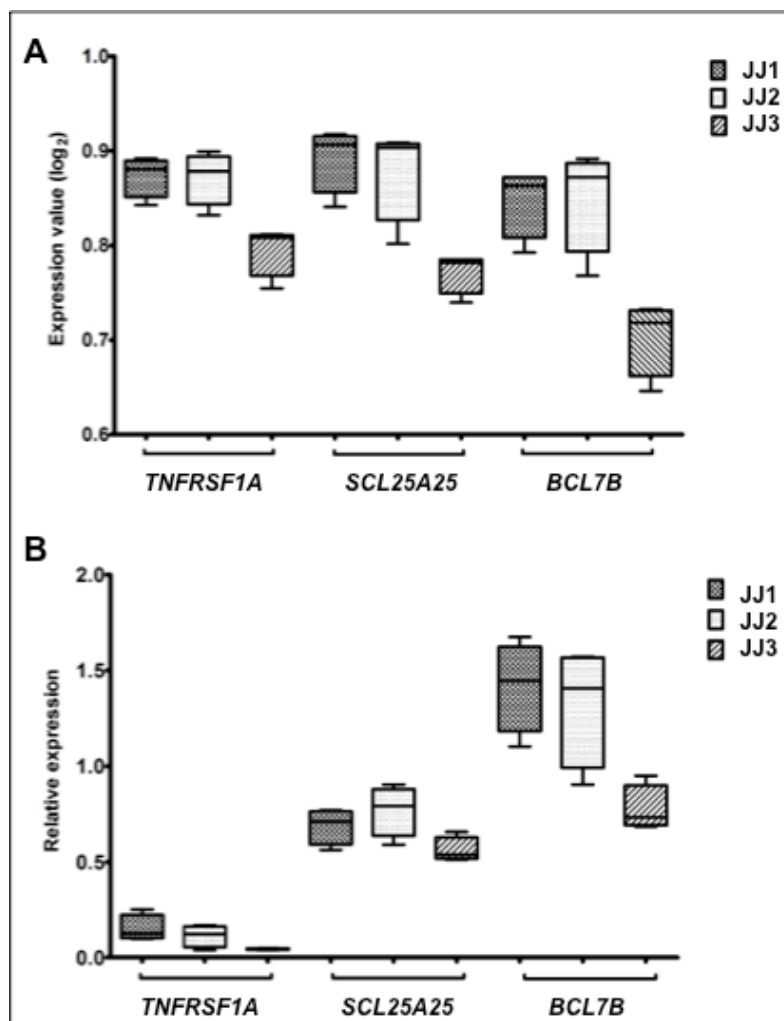


**Figura 28** - Curva de amplificação dos genes endógenos candidatos a genes de referência por qPCR. **A** - *RPLP0* e *PPIA* para o grupo Caco-2/EH41. **B** - *MRPL19* e *TBP* para o grupo Caco-2/Ec472

As Figuras 29 e 30 mostram os resultados de qPCR para os genes *TBKBP1*, *PRIMA1* e *NEGR1* do grupo Caco-2/EH41 e para os genes *TNFRSF1A*, *SLC25A25* e *BCL7B* em Caco-2/Ec472. A análise comparativa dos padrões de variação dos boxplots obtidos para os dados de expressão provenientes de microarranjos de DNA e dos dados de expressão obtidos por qPCR revelam padrões de variação similares. Estes resultados de qPCR validam, portanto, os dados obtidos por microarranjos de DNA.



**Figura 29** - Validação dos dados de expressão obtidos por microarranjos de DNA do grupo Caco-2/EH41. **A** - Gráfico boxplot dos valores de expressão obtidos por microarranjos de DNA dos genes *TBKBP1*, *PRIMA1* e *NEGR1* para as amostras de cada janela justaposta (JJ). **B** - Gráfico boxplot dos valores de expressão relativa obtidos por qPCR dos genes *TBKBP1*, *PRIMA1* e *NEGR1* para as amostras de cada janela



**Figura 30** - Validação dos dados de expressão obtidos por microarranjos de DNA do grupo Caco-2/Ec472. **A** - Gráfico boxplot dos valores de expressão obtidos por microarranjos de DNA dos genes *TNFRSF1A*, *SLC25A25* e *BCL7B* para as amostras de cada janela. **B** - Gráfico boxplot dos valores de expressão relativa obtidos por qPCR dos genes *TNFRSF1A*, *SLC25A25* e *BCL7B* para as amostras de cada janela

## 5 DISCUSSÃO

As STECs são importantes patógenos humanos por causarem desde diarreias até a SHU. No caso da SHU, há diversos sorotipos associados a essa síndrome, tais como O157:H7 e O113:H21 (Kaper et al., 2004; Bugarel et al., 2010; Monaghan et al., 2011; Buvens; Pierard, 2012). No Brasil, entretanto, o sorotipo O113:H21 ainda não aparece associado a SHU, embora seja frequentemente isolado de carcaças e fezes bovinas (dos Santos et al., 2010; Gonzalez et al., 2016). Cabe aqui notar que alguns isolados de O113:H21 no Brasil, bem como isolados de pacientes com SHU na Argentina, pertencem a um mesmo grupo clonal (Feng et al., 2014). Assim, muitos estudos têm sido realizados em isolados de STEC provenientes de animais e do meio ambiente visando a identificar genes de virulência relacionados com SHU, além de caracterizar molecular ou fenotipicamente esses isolados para determinar o seu potencial patogênico (Feng et al., 2014; Franz et al., 2015; Feng et al., 2017).

Nosso grupo já investigou comparativamente o perfil transcricional e as redes de coexpressão gênica de STEC EH41 (associado à SHU) e Ec472/01 (isolado de fezes bovina) (Bando et al., 2017). A análise comparativa do perfil transcricional de EH41 e Ec472/01 revelou que os transcritos diferencialmente expressos estão hipoexpressos em EH41 e que a maioria destes genes estão envolvidos em processos metabólicos. Analisando-se o crescimento destes isolados verificou-se que EH41 cresce significativamente menos que Ec472/01. Mais ainda, foi constatado que somente EH41 expressa um conjunto de genes - nesse trabalho denominados genes EE - que inclui o regulador transcricional *dicA*. Verificou-se ainda, por meio de PCR, que *dicA* não está presente em outros isolados de O113:H21 provenientes de animais e do meio ambiente.

A rede de coexpressão gênica (RCG) mostrou que a maioria de genes EE está situada em um único módulo transcricional, que foi denominado módulo *dicA*. Curiosamente, seus componentes são, em sua maioria, derivados de profagos. Sabe-se que profagos podem fornecer muitas vantagens adaptativas,

como resistência a antibióticos, tolerância ao estresse oxidativo e/ou ácido, permitindo a bactéria sobreviver em meio ambiente hostil (Wang et al., 2010).

A proteína DicA em *E. coli* é um repressor transcricional sensível a temperatura e atua no crescimento bacteriano (Bejar et al., 1986; Bejar et al., 1988). Além disso, DicA apresenta similaridade significativa na sequência de amino ácidos com RovA e SlyA, que são reguladores de genes de virulência em *Yersinia* e *Salmonella*, respectivamente (Yun et al., 2012). Estas duas proteínas regulam vários processos fisiológicos bacterianos envolvidos na sobrevivência e adaptação ao estresse ambiental e na virulência (Daniels et al., 1996; Tran et al., 2005; Ellison; Miller, 2006). Em *E. coli* SlyA induz a expressão do gene *hlyE*, que codifica para a toxina hemolisina E (Spory et al., 2002).

A análise de RCG de EH41 indicou que *dicA* regula positivamente os genes localizados no módulo *dicA* (possivelmente são genes de virulência) e pode estar regulando negativamente os genes envolvidos com a crescimento bacteriano. Estudos mostraram que pacientes apresentavam positividade para STEC mesmo após nove meses sem sintomas de doença (Matussek et al., 2016). Assim, é possível que *dicA* esteja envolvido na adaptação de EH41 para sobrevivência por longos períodos no meio ambiente entérico e, conseqüentemente, causar doença grave, como SHU em humanos.

Devido a estas diferenças das duas STECs, um estudo da resposta das células Caco-2 após três horas de interação com EH41 e Ec472/01 foi investigado através da análise de RCG. A rede do grupo Caco-2/EH41 apresentou muitas ligações gene-gene negativas e positivas. Isto indica uma desregulação da expressão gênica global. Já nas redes do grupo Caco-2/Ec472 ou de Caco-2 controle (sem exposição bacteriana) a maioria dos genes possuem ligações positivas e poucos genes - que estão na borda do módulo transcricional ou conectando dois módulos - possuem ligações negativas. Este tipo de rede é o esperado de uma célula "normal", isto é, nós com ligações positivas tendem a formar um módulo, enquanto que nós com ligações negativas normalmente atuam como conectores entre módulos (Barabási; Oltvai, 2004; Liu et al., 2012; Lin et al., 2013). Estes resultados sugerem que EH41 é capaz de induzir uma intensa desregulação genômica no enterócito.

Mais ainda, a maioria dos hubs da rede do grupo Caco-2/EH41 está envolvida na resposta inflamatória e imune. Por outro lado, somente um terço dos hubs da rede do grupo Caco-2/Ec472 está envolvido na imunidade inata. Existe evidência de que a inflamação aguda tem um papel importante no desenvolvimento de SHU. Pacientes com SHU apresentaram neutrofilia, aumento da proteína c reativa e também das citocinas pró-inflamatórias, indicando que a colite hemorrágica pode estar associada com a evolução de complicações severas, como a SHU e encefalopatias (Shimizu et al., 2012).

Portanto, este trabalho centrou-se numa abordagem de biologia de sistemas onde se integrou a análise genômica e fenotípica de Caco-2 durante a interação enterócito-bactéria. O estudo genômico baseou-se no estudo temporal de redes de coexpressão gênica para compreender os mecanismos moleculares que estão envolvidos na patogenicidade de STEC associada (EH41) e não associada (Ec472/01) à SHU. As alterações fenotípicas ocorridas nas células Caco-2 ao longo da exposição a cada um dos isolados de STEC foram visualizadas através de MEV.

A análise genômica, como mostrado na seção anterior, visou à investigação da variação topológica das RCG das células Caco-2 ao longo de três horas de interação com EH41 (grupo Caco-2/EH41) ou Ec472/01 (grupo Caco-2/Ec472) (Figura 9). A interação com a bactéria causa uma perturbação - perda do status *scale free* - nas RCG dos dois grupos a partir de uma hora de interação enterócito-bactéria. É possível notar, pela visualização de redes 3D, uma alteração mais pronunciada nas redes JJ3 a JJ6 do grupo Caco-2/EH41, quando comparadas com as redes da Caco-2/Ec472 nesse mesmo período. Após uma hora e meia de interação enterócito-bactéria, observou-se a reorganização (restauração do status *scale-free*) da rede de Caco-2/EH41. Por outro lado, na rede de Caco-2/Ec472, essa reorganização não chegou a ocorrer.

Muitos trabalhos tem mostrado que modificações topológicas nas RCG constituem a força motriz para alterações no metabolismo celular, caracterizando a mudança de um estado de equilíbrio para outro sob influência de fatores endógenos ou exógenos (Sieberts; Schadt, 2007; Sahni et al., 2013). Isso é particularmente bem caracterizado nas transições saúde-doença (Liu et



al., 2011). Ao se investigar a topologia das redes dos grupos Caco-2/EH41 e Caco-2/Ec472 foi possível identificar as diferentes etapas de uma transição saúde-doença: um estado normal, um estado de pré-doença (com a perda do status *scale free*), e o estado de doença (nova topologia da rede). Mais ainda, as imagens de MEV mostraram a dinâmica das alterações morfológicas do enterócito durante contato com STEC: enterócitos em contato com EH41 apresentaram maior perda de microvilosidades do que aqueles em contato com Ec472/01.

Considerando que modificações topológicas levam a alterações da estrutura modular das RCG, grande parte deste trabalho centrou-se na identificação e caracterização de módulos transcricionais, com o objetivo de estudar a dinâmica das vias moleculares envolvidas na resposta diferencial do enterócito a EH41 ou Ec472/01.

### 5.1 Caco-2/EH41

Utilizando-se da ferramenta WGCNA foi possível obter uma rede *scale free* e assim correlacionar módulos transcricionais com as janelas ao longo de três horas de interação enterócito-bactéria. Vale a pena mencionar que a rede do grupo Caco-2/EH41 está dividida em dois meta-módulos (Figura 13A), onde cada um deles está relacionado com uma janela. Este resultado indica que durante a resposta da Caco-2 à EH41 há uma expressão gênica diferencial significativa nos diferentes intervalos ao longo de três horas de interação enterócito-bactéria. Desta forma foi possível identificar cinco módulos transcricionais positivamente associados a cada janela: i) módulo *blue* para JJ1; ii) módulos *turquoise* e *pink* para JJ2 e; iii) módulos *brown* e *yellow* para JJ3 (Figura 13B).

Investigando-se as funções biológicas dos hubs destes módulos foi possível correlacionar a dinâmica das vias moleculares com a resposta do enterócito ao longo de três horas de interação enterócito-bactéria. Na primeira hora há poucos hubs envolvidos na resposta imune inata, entretanto a partir da segunda hora muitos hubs estão envolvidos na resposta inflamatória e imune, assim como na apoptose. É interessante notar que é na janela JJ2 (período de 1 a 2 horas) que muitos genes e hubs estão envolvidos na resposta imune inata, principalmente na resposta inflamatória (Figuras 15 e 17A, Tabela 3).

Três hubs associados à JJ1 merecem destaque: *ADRA1A*, *TBKBP1* e *CSH1*. A proteína codificada pelo hub *ADRA1A* está envolvida na regulação positiva da cascata de sinalização de ERK1/2 (*Extracellular signal-Regulated Kinase*) e MAPK (*Mitogen Activated Protein Kinase*). Sabe-se que a ativação de ERK1/2 é induzida por Stx e contribui para a expressão de citocinas pró-inflamatórias no epitélio intestinal (Jandhyala et al., 2016).

O segundo hub *TBKBP1* codifica para uma proteína adaptadora que se liga ao TBK1 e faz parte da via TNF/NFK $\beta$  (Bouwmeester et al., 2004). *TBKBP1* ainda está envolvido em outras vias de sinalização na resposta imune inata: por exemplo, como regulador na indução de IFN1 (Goncalves et al., 2011), sendo também importante na autofagia e no câncer (Helgason et al., 2013).

O terceiro hub *CSH1*, *chorionic somatomammotropin hormone 1*, codifica para uma proteína da família de hormônios de crescimento, mas também pode atuar na cascata de sinalização JAK-STAT. Segundo Varelle et al. (2008), as EHEC diminuem a ativação do STAT-1 pela redução da expressão de mRNA de iNOS induzida por IFN- $\gamma$ , assim como diminui a produção de óxido nítrico (NO) em células Caco-2. NO é capaz de modular o processo inflamatório e causar a morte de patógenos. As fontes químicas ou celulares de NO inibem a expressão do gene *stx2* e a sua síntese sem afetar o crescimento e sobrevivência da bactéria. Portanto, a limitação da produção de NO pode ser uma estratégia adotada por EHEC que favorece sua virulência.

Na janela JJ2 mais da metade dos hubs (14 de 21) estão envolvidos na resposta imune inata, sendo a resposta inflamatória - que envolve sete hubs - o processo mais importante. Os hubs *TICAM1* (aliase *TRIF*), *BCL3*, *TNFRSF1A* e *BIRC3* codificam para proteínas envolvidas na regulação de IKK (I-kappa $\beta$  kinase) e sinalização de NF-k $\beta$  (Nuclear factor kappa  $\beta$ ). Berin et al. (2002) descreveram que a infecção por *E. coli* enterohemorrágica (EHEC) ativa a sinalização de NF-k $\beta$  e MAP kinase (p30 e ERK 1/2). Essas vias são requeridas na indução de uma resposta pró-inflamatória das células epiteliais intestinais. Mais ainda, a via dependente de TRIF ativa o *interferon regulatory factor 3* (IRF3), além de NF-k $\beta$  e AP-1, que induzem a expressão de interferon (IFN) tipo I (Lee; Kim, 2007). *BCL3* também interage com o complexo AP-1

mostrando efeitos pró-inflamatório e proliferativo em resposta à UPEC nas células uroteliais (Mysorekar et al., 2002).

Sabe-se que a proteína TRIF, no epitélio intestinal, está envolvida na comunicação patógeno-hospedeiro via TLR5 que reconhece as fímbrias bacterianas e induz a resposta inflamatória (Lee; Kim, 2007; Choi et al., 2010). É interessante mencionar que, neste intervalo de tempo, EH41 apresenta fímbrias (imagem T15, Figura 23). Este resultado indica que Caco-2 reconhece, via TLR5, a fímbria expressa pela EH41, ativando em seguida citocinas pró-inflamatórias. Este cenário não foi observado no grupo Caco-2/Ec472 e pode estar associado à ausência de fímbrias em Ec472/01 (Figura 23).

Outros três hubs, *IRF1*, *CLIC1* e *ADRBK1* (aliase *GRK2*) também codificam proteínas envolvidas no processo pró-inflamatório. IRF1 por meio de um mecanismo elaborado e cooperativo induz a produção de IL-7 essencial na regulação da resposta imune na mucosa intestinal humana (Oshima et al., 2004). A proteína ADRBK1 está envolvida na sinalização de TNF- $\alpha$  em células epiteliais do cólon e associado a inflamação intestinal (Steury et al., 2017). A proteína CLIC1 está envolvida na sinalização para a ativação de IL-1 $\beta$  e de NLRP3 (Domingo-Fernández et al., 2017). O inflamassoma NLRP3 regula a ativação de caspase-1 que leva a maturação das citocinas pró-inflamatórias IL-1 $\beta$  e IL-18.

Outro processo biológico relevante nesta janela (JJ2) é a apoptose (seis hubs). Sabe-se que a apoptose é um dos processos induzidos por Stxs (Tang et al., 2015). Os hubs *BBC3*, *UFC1*, *AKAP8* (aliase *AKAP95*), *BCL2L4*, *IFI27* (aliase *ISG12*) e *NPIP* codificam proteínas envolvidas na apoptose. Vários trabalhos mostraram que proteínas codificadas por alguns desses genes estão envolvidas na apoptose de enterócitos, no câncer de células colorretais ou são ativadas por lipopolissacarídeos bacterianos (LPS), como descrito a seguir.

A proteína BBC3 (aliase *PUMA*) está envolvida na apoptose das células epiteliais intestinais e pode contribuir para a patogênese da colite ulcerativa. (Qiu et al., 2011). Já o AKAP95 está envolvido na supressão da produção de TNF- $\alpha$  induzida por LPS (Wall et al., 2009). Yu et al. (2016) mostraram que o *long intergenic noncoding UFC1* (linc-*UFC1*) promove a proliferação celular enquanto atenua a apoptose em câncer colorretal. Dessa forma participa da

regulação do ciclo celular, da apoptose intrínseca e da sinalização de  $\beta$ -catenina e P38 (Yu et al., 2016; Jandhyala et al., 2016).

Ainda em JJ2, um hub, *DOK1*, parece atuar como modulador da resposta imune. Este gene codifica para um inibidor dos percursores de sinalização RAS/ERK e PI3K/Akt, atuando, portanto, como um regulador negativo das vias de sinalização para ativação de células NK humanas (Celis-Gutierrez et al., 2014).

Finalmente, na janela JJ3 há menos hubs envolvidos na resposta imune inata: somente nove de 33 hubs. Quatro hubs estão envolvidos na resposta inflamatória: *BAHD1*, *CEACAM20*, *CXCL5* e *RPS27A*. Três hubs estão envolvidos na resposta imune: *HIST1H2BJ*, *RAB4A* e *SIRPB1*. Há também dois hubs, *NUP62* e *PARL*, envolvidos no processo anti-apoptótico.

As funções biológicas de alguns desses hubs serão comentadas a seguir. O hub *RAB4A*, que também aparece como hub na rede do grupo Caco-2/Ec472, codifica um membro da família Ras que participa da transdução de sinal das proteínas Rab, além da apresentação e processamento de antígeno (Goueli et al., 2012). Outro hub importante é *CXCL5* (alias *ENA-78*), que codifica para uma quimiocina e pode ser induzido por stx1 (Thorpe et al., 2001). O hub *BAHD* codifica para um regulador da via de sinalização de TNF. A proteína BAHD está hipoexpressa nas células do cólon de pacientes com colite ulcerativa (Zhu et al., 2015). O hub *CEACAM20* codifica para uma proteína da família de Ig transmembrana específica de microvilos de enterócitos que se liga a tirosina kinase ativando NF- $\kappa$ B, além de induzir a produção de citocinas, como a IL-8 (Murata et al., 2015).

Dois outros hubs estão envolvidos na regulação negativa da via de sinalização apoptótica. *PARL* é um regulador negativo da via de sinalização apoptótica intrínseca. Segundo Lee et al. (2010), as vias indutoras de apoptose intrínseca e/ou extrínseca estão envolvidas no processo de morte celular após intoxicação. *NUP62* participa da regulação negativa do processo apoptótico e morte celular programada.

Outro aspecto importante do presente trabalho foi a análise fenotípica das células Caco-2 durante a interação com STEC. Foram observadas alterações na morfologia das microvilosidades da célula Caco-2 (Figura 23), como a perda

de microvilosidades (imagem T5) a partir de 45 min após interação com EH41 e esta perda se intensificou após este intervalo (imagens T15 e T24). É interessante citar que três hubs - *ROPN1L*, *MARCKS* e *OCN* - codificam proteínas envolvidas com o citoesqueleto celular: o movimento ciliar do epitélio (Grizzi et al., 2015), o componente do filamento de actina (Hartwig et al., 1992) e a organização da junção celular (Lapointe; Buret, 2012), respectivamente. Um destes hubs, *MARCKS*, pode estar diretamente envolvido nestas alterações fenotípicas, por codificar para uma proteína envolvida na regulação do citoesqueleto da actina (Hartwig et al., 1992). Mais ainda, *MARCKS* pode estar envolvido na expressão de citocinas pró-inflamatórias em macrófagos após indução por LPS (Lee et al., 2015).

## 5.2 Caco-2/Ec472

A análise de clusterização dos módulos eigengenes e das janelas da rede do grupo Caco-2/Ec472 revelou menor correlação dos meta-módulos com as janelas (Figura 18A). Esse resultado difere do que foi encontrado para o grupo Caco-2/EH41 e indica que o perfil transcricional entre as janelas apresenta menor variação. Mesmo assim, foi possível obter cinco módulos correlacionados positivamente com as janelas temporais: i) módulo *yellow* para JJ1; ii) módulo *blue* para JJ2 e; iii) módulos *turquoise*, *brown* e *pink* para JJ3 (Figura 18B).

O enriquecimento funcional (Figura 20) mostrou que relativamente poucos genes estão envolvidos na resposta imune inata até duas horas de interação enterócito-bactéria. Porém, após esse período, uma quantidade significativa de genes parece estar envolvida na resposta imune, principalmente no módulo *turquoise*.

Comparativamente ao grupo Caco-2/EH41, o grupo Caco-2/Ec472 apresentou poucos hubs que codificam proteínas envolvidas na resposta imune: somente 13 de 59 hubs (Figura 22B). Outras diferenças também devem ser mencionadas, tais como: i) ausência de hubs envolvidos na resposta inflamatória. ii) dois hubs *PCYT2* e *ACTR2* (aliase *ARP2*) envolvidos em fagocitose (Figura 22A) (Bakovic et al., 2007; May et al., 2000; Wiedemann et al., 2001). Sabe-se que a fagocitose é um processo importante para a

manutenção da homeostase celular e também possui um papel crucial na defesa contra patógenos (Henneke; Golenbock, 2004, Winchell et al., 2016).

Na janela JJ1, somente um hub parece estar relacionado com inflamação. O gene *KCNN4* que codifica um dos componentes de canais de potássio e está envolvido na migração de enterócitos. Zundler et al. (2016) sugerem que disfunção na migração de células epiteliais intestinais podem contribuir para inflamação crônica intestinal. Outro hub, *E2F2*, codifica para um fator de transcrição e está envolvido no processo de anti-apoptose (Chen et al., 2015; Chambers et al., 2015). É curioso que três hubs - *CHAF1A*, *RAD51* e *UBE2T* - estão envolvidos no reparo de DNA (Yu et al., 2015; Bhattacharya et al., 2017; Kelsall et al., 2012). Bielaszewska et al. (2017) mostraram que algumas toxinas, como a toxina citoletal CdtV-B de EHEC O157, provoca danos no DNA celular ativando respostas para o reparo de DNA.

Na janela JJ2, é possível que três hubs estejam envolvidos com a resposta imune: *RAB4A*, *NEURL* e *ACTR2*. O primeiro hub, também hub na rede do grupo Caco-2/EH1, está envolvido na apresentação e processamento de antígenos, o segundo hub codifica para uma proteína envolvida na ubiquitinação e apoptose. Já o terceiro hub, *ACTR2*, codifica para uma proteína estrutural do citoesqueleto (ARP2). Mais ainda, alguns estudos mostraram que o complexo ARP2/3 está envolvido na formação do fagossomo (May et al., 2000; Wiedemann et al., 2001). Outro hub interessante desta janela é *SURF4*, que codifica para uma proteína de membrana envolvida no transporte de proteínas do complexo de Golgi para o retículo endoplasmático. Esse tipo de transporte é utilizado para a internalização de Stx (Bielaszewska et al., 2017).

Quanto à janela JJ3 (três horas de interação enterócito-bactéria), obtiveram-se três módulos associados a esse intervalo, o que indica um início de reorganização hierárquica da RCG. Nessa janela, o módulo *turquoise* parece estar mais relacionado com a resposta imune: seis de 11 hubs estão relacionados a essa função biológica (Figura 22A). Diferentemente do grupo Caco-2/EH41, três destes hubs - *UPF1*, *HAX1* e *TRAFD1* - do módulo *turquoise* estão envolvidos na modulação da resposta imune, mantendo assim a homeostase do sistema imune. A função das proteínas codificadas por esses genes estão descritas a seguir.

A proteína UPF1, que atua juntamente com Regnase-1 e pode ser induzida por LPS, restringe a inflamação e mantém a homeostase da resposta imune. (Mao et al., 2017). A proteína multifuncional codificada pelo gene *HAX1* está associada ao estímulo por IL-1 e envolvida na regulação negativa de apoptose (Zayat et al., 2015). *TRAFD1* atua como regulador negativo da resposta imune inata excessiva contra infecção microbiana através das vias de sinalização TLR4 e RIG-I-like helicase (Sanada et al., 2008).

A análise fenotípica das células Caco-2 durante a interação com Ec472/01 mostrou alterações na morfologia das microvilosidades do enterócito (Figura 23). Entretanto, essas alterações fenotípicas diferem das observadas para células Caco-2 em contato com EH41: a perda de microvilosidades é menor e ocorre somente após duas horas de interação enterócito-bactéria (imagem T15). Mais ainda, os microvilos de enterócitos em contato com Ec472/01 estão mais alongados, quando comparados aos dos enterócitos expostos à EH41.

Na rede do grupo Caco-2/Ec472 seis hubs estão envolvidos com o citoesqueleto celular. Quatro desses hubs - *PODXL*, *EMP2*, *POF1B* e *NAPA* - codificam proteínas do componente estrutural do filamento de actina ou estão envolvidos na organização do citoesqueleto (Naishiro et al., 2005; Morales et al., 2009; Crespi et al., 2015; Lechuga et al., 2017). Os dois outros hubs, *CCNA2* e *PPP1R15B*, estão envolvidos na regulação da reorganização do citoesqueleto (Loukil et al., 2015; Chambers et al., 2015).

Neste trabalho foi possível analisar comparativamente, por análise genômica, a dinâmica da resposta das células Caco-2 durante a interação com EH41 e Ec472/01. Já na primeira análise da variação topológica das RCG (Figura 9) é possível observar que a dinâmica de expressão global dos genes é diferente, determinando assim as diferentes topologias das redes nos dois grupos.

No início da interação enterócito-bactéria as redes dos grupos Caco-2/EH41 e Caco-2/Ec472 possuem status *scale free*. Com a progressão do processo de interação enterócito-bactéria, observa-se que no grupo Caco-2/EH41 as redes rapidamente perdem esse status, readquirindo essa condição somente ao final do período de interação aqui analisado. Já nas redes do grupo

Caco-2/Ec472 a perda do status *scale free* ocorre mais lentamente (Figura 9), porém esse status não é recuperado até o término do experimento.

A análise de módulos transcricionais mostrou que durante a interação enterócito-bactéria diferentes módulos eigengenes estão associados a três intervalos de tempo. Esses módulos foram funcionalmente caracterizados quanto aos processos biológicos envolvidos na resposta ao estímulo bacteriano. Através dessa análise foi possível compreender dinamicamente os mecanismos moleculares envolvidos nessa resposta. Mais ainda, tornou-se evidente que existem diferenças relevantes no padrão de resposta das células Caco-2 durante a interação com EH41 ou com Ec472/01, como discutido a seguir.

A principal diferença está na indução da resposta inflamatória. No grupo Caco-2/EH41, por exemplo, há muitos hubs envolvidos na cascata de sinalização MAPK, ERK1/2, JAK-STAT e NF- $\kappa$ B durante todo o período de interação enterócito-bactéria. Já no grupo Caco-2/Ec472 este processo não é ativado ao longo desse período de interação (Figuras 17 e 22). Existem evidências de que a inflamação aguda está associada ao desenvolvimento de SHU. Pacientes com essa síndrome apresentaram um aumento na proteína c reativa, neutrofilia e uma elevada concentração de citocinas pró-inflamatórias circulantes. Assim, a colite hemorrágica pode conduzir a uma evolução para doenças mais graves, como a SHU e a encefalopatia (Shimizu et al., 2012).

Outra distinção entre os dois grupos é o processo de homeostase do sistema imune. Somente no grupo Caco-2/Ec472 há hubs (janela JJ3) que estão envolvidos na modulação da resposta inflamatória excessiva.

Hubs envolvidos na apoptose estão presentes em ambos os grupos. Entretanto, no grupo Caco-2/EH41 há nove hubs relacionados a essa função biológica, sendo que seis deles estão nos módulos da janela JJ2. Comparativamente, no grupo Caco-2/Ec472 somente dois hubs estão envolvidos com apoptose.

Ambos os grupos possuem hubs nas janelas JJ1 e JJ3 envolvidos na organização do citoesqueleto. As imagens de MEV mostraram, nos dois grupos, alteração da morfologia da bordadura celular de Caco-2 já na primeira hora de interação enterócito-bactéria. Entretanto, existe uma destruição das



microvilosidades mais evidente nas células em contato com EH41, em comparação com os enterócitos em contato com Ec472/01 (Figura 23). Sabe-se que microvilos servem como uma barreira eletrostática para a adesão bacteriana (Bennet et al., 2014). Desta forma, a destruição dos microvilos que é observada nas células Caco-2 em contato com EH41 pode favorecer a adesão bacteriana no ambiente entérico.

## 6 CONCLUSÕES

O presente trabalho centrou-se na análise dinâmica das RCG das células Caco-2 durante a interação com dois isolados de STEC do sorotipo O113:H21: EH41, isolado de um paciente com SHU, e Ec472/01, isolado de fezes bovina.

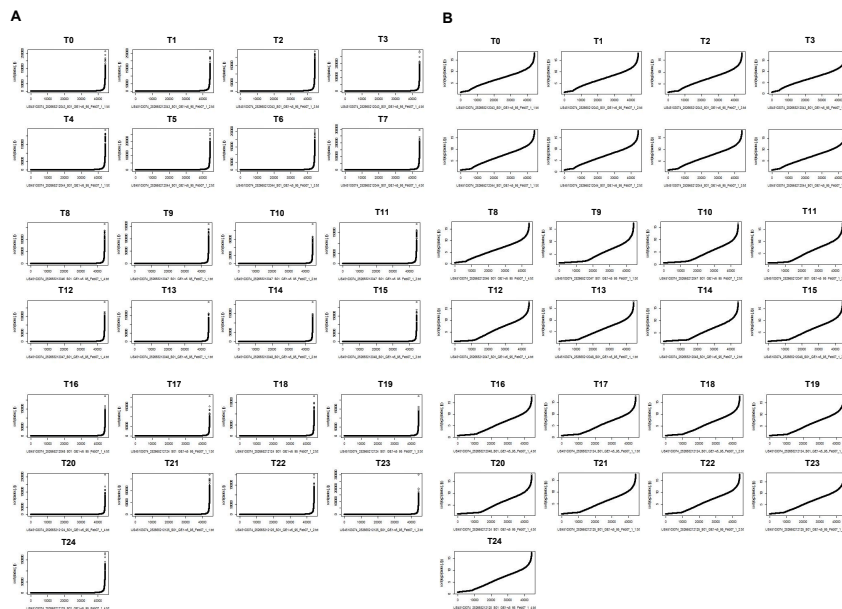
A análise genômica mostrou que o mecanismo molecular da resposta de Caco-2 durante a interação com EH41 ou Ec472/01 é claramente distinto. Nas redes do grupo Caco-2/EH41 as alterações topológicas incluíram a perda do status *scale free* e a sua recuperação, com o estabelecimento de uma nova hierarquia de genes na rede. Esses resultados se enquadram no modelo de transição saúde-doença proposto por Liu et al. (2011): a nova rede representa a resposta adaptativa da célula ao patógeno, o que não significa um retorno à normalidade (ver adiante). Já no grupo Caco-2/Ec472 as redes, após a perda do status *scale free*, não recuperam esse status até o final do período estudado, o que sugere um estado de transição mais prolongado para reorganização da hierarquia da rede.

Mais ainda, através da caracterização dos módulos transcricionais, foi possível compreender dinamicamente os mecanismos moleculares envolvidos na resposta diferencial de Caco-2 aos dois isolados aqui estudados. STEC EH41 induz rapidamente a resposta inflamatória e apoptótica a partir da primeira hora de interação enterócito-bactéria. Por outro lado, as células Caco-2 em contato com Ec472/01 ativam, a partir de uma hora, mecanismos de defesa contra patógenos invasores, como a fagocitose e, a partir da segunda hora, expressam moduladores da homeostase imune.

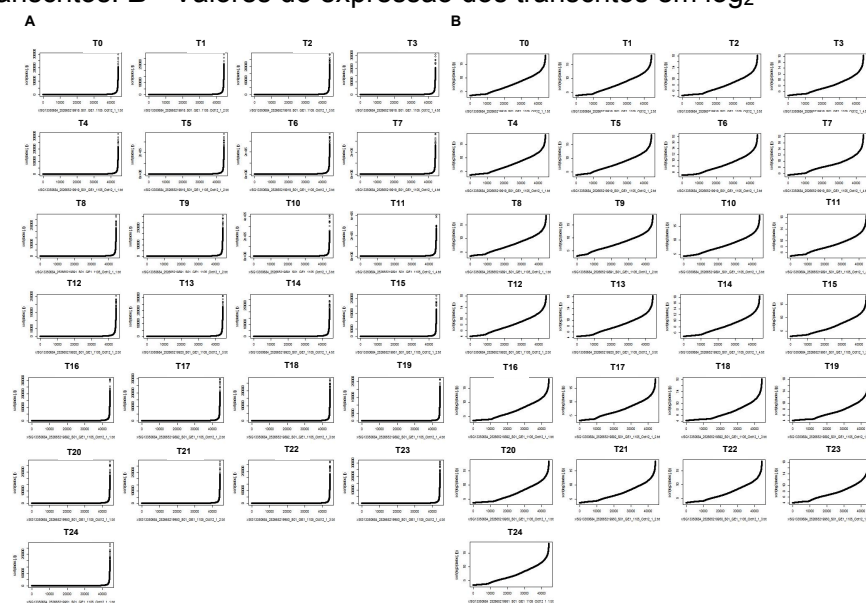
A análise fenotípica das células Caco-2 mostrou, de forma nítida, uma maior destruição dos microvilos dos enterócitos em contato com EH41 do que com Ec472/01. Integrando os resultados genômicos e fenotípicos pode-se concluir que EH41 induz em Caco-2 - em comparação com Ec472/01 - maiores e mais rápidas alterações na expressão gênica global, além de uma resposta inflamatória e apoptótica excessiva, levando assim a alterações morfológicas mais pronunciadas nas células Caco-2.

Em seu conjunto, esses resultados contribuem para uma melhor compreensão dos mecanismos moleculares envolvidos na patogenicidade das STECs associadas à SHU. Assim, as perspectivas de desenvolvimento deste trabalho deverão incluir a investigação de fatores de virulência e vias moleculares envolvidas na indução das respostas imunes que podem conduzir à SHU.

## 7 ANEXOS



**Anexo 1** - Análise de qualidade da hibridização nos microarranjos de DNA de cada amostra do grupo Caco-2/EH41. **A** - Valores de expressão dos transcritos. **B** - Valores de expressão dos transcritos em log<sub>2</sub>



**Anexo 2** - Análise de qualidade da hibridização nos microarranjos de DNA de cada amostra do grupo Caco-2/Ec472. **A** - Valores de expressão dos transcritos. **B** - Valores de expressão dos transcritos em log<sub>2</sub>

### Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41

Módulo	KEGG	valor-p	Genes
Blue	Salivary secretion	0.005	PRB1;ATP1A3;PRH2;ATP2B3;ITPR3;CALML3;ADRB1;AQP5;ADRA1A;SLC9A1
	Calcium signaling pathway	0.005	CAMK2D;NOS3;PTAFR;TACR2;CACNA1A;ATP2B3;CALML3;ITPR3;ADRB1;CACNA1C;ADRA1A;CACNA1I;STIM1;P2RX2;PLCG2;BDKRB2
	cAMP signaling pathway	0.006	MAP2K1;CAMK2D;PPP1R12A;BAD;GIPR;GPR119;GNAI3;ATP1A3;ATP2B3;CALML3;ADRB1;CACNA1C;SLC9A1;GRIN3A;RRAS;PDE4A;CNGB1
	Phagosome	0.007	DYNC2H1;DYNC1I2;NCF1;ATP6AP1;SFTP2;SFTPA2;SFTPD;TAP2;HLA-A;HLA-G;THBS1;HLA-DRB3;HLA-DRB1;VAMP3;ATP6V1F
	Proteoglycans in cancer	0.007	WNT10B;MAP2K1;CAMK2D;TGFB1;PPP1R12A;WNT3A;FZD7;FZD6;ITPR3;THBS1;SLC9A1;MAPK11;RRAS;PLCG2;SDC1;FLNB;MET
	Renal cell carcinoma	0.007	EGLN1;MAP2K1;TGFB1;PDGFB;TGFA;TCEB1;MET;CRKL
	Pathways in cancer	0.007	SPI1;CSF3R;PDGFB;GNAI3;TGFA;RASGRP4;CRKL;CKS1B;IKBKB;PLCG2;BDKRB2;GNG8;E2F2;TCEB1;JAK1;EGLN1;WNT10B;MAP2K1;TGFB1;BAD;WNT3A;FZD7;FZD6;COL4A2;MSH2;RAR;BIRC7;MET
	Autoimmune thyroid disease	0.007	TPO;GZMB;HLA-A;HLA-DRB3;HLA-G;HLA-DRB1;TSHB
	Graft-versus-host disease	0.008	GZMB;HLA-A;KIR3DL2;HLA-DRB3;HLA-G;HLA-DRB1
	HTLV-I infection	0.008	ATF1;WNT10B;SPI1;TGFB1;NFYB;WNT3A;FZD7;FZD6;PDGFB;HLA-A;ITGAL;HLA-G;IKBKB;TERT;RRAS;E2F2;TCF3;HLA-DRB3;HLA-DRB1;JAK1
	cGMP-PKG signaling pathway	0.013	MAP2K1;PPP1R12A;BAD;NOS3;GNAI3;ATP1A3;ATP2B3;CALML3;ITPR3;ADRB1;CACNA1C;ADRA1A;BDKRB2;CNGB1

continua

**Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41** *continuação*

<i>Epstein-Barr virus infection</i>	0.014	HDAC5;SPI1;HLA-A;ITGAL;HLA-G;IKBKB;CCNA2;MAPK11;IRAK1;PLCG2;POLR1D;POLR2E;HLA-DRB3;HLA-DRB1;POLR2K;JAK1
<i>Neurotrophin signaling pathway</i>	0.014	IKBKB;MAPK11;MAP2K1;CAMK2D;IRAK1;BAD;KIDINS220;NTF3;PLCG2;CALML3;CRKL
<i>Proximal tubule bicarbonate reclamation</i>	0.016	MDH1;ATP1A3;SLC38A3;GLS
<i>Antigen processing and presentation</i>	0.017	CD4;NFYB;TAP2;HLA-A;KIR3DL2;HLA-DRB3;HLA-G;HLA-DRB1
<i>Malaria</i>	0.018	TGFB1;CD81;SDC1;ITGAL;MET;THBS1
<i>Glycosphingolipid biosynthesis - globo series</i>	0.020	B3GALNT1;NAGA;A4GALT
<i>Glioma</i>	0.021	MAP2K1;CAMK2D;PDGFB;PLCG2;TGFA;E2F2;CALML3
<i>Cholinergic synapse</i>	0.021	ACHE;MAP2K1;CAMK2D;KCNQ3;GNAI3;GNG8;CACNA1A;ITPR3;FYN;CACNA1C
<i>Allograft rejection</i>	0.022	GZMB;HLA-A;HLA-DRB3;HLA-G;HLA-DRB1
<i>Pancreatic cancer</i>	0.023	IKBKB;MAP2K1;TGFB1;BAD;TGFA;E2F2;JAK1
<i>Lysine degradation</i>	0.023	GCDH;ECHS1;SUV39H1;NSD1;SETD1A;PIPOX
<i>Toxoplasmosis</i>	0.031	IKBKB;MAPK11;TGFB1;IRAK1;BAD;GNAI3;BIRC7;HLA-DRB3;HLA-DRB1;JAK1
<i>Chagas disease (American trypanosomiasis)</i>	0.035	IKBKB;MAPK11;TGFB1;ACE;IRAK1;SERPINE1;BDKRB2;GNAI3;CFLAR
<i>Type I diabetes mellitus</i>	0.036	GZMB;HLA-A;HLA-DRB3;HLA-G;HLA-DRB1
<i>Leishmaniasis</i>	0.037	MAPK11;TGFB1;NCF1;IRAK1;HLA-DRB3;HLA-DRB1;JAK1
<i>Platelet activation</i>	0.037	COL1A1;MAPK11;PPP1R12A;STIM1;NOS3;TBXAS1;PLCG2;GNAI3;ITPR3;FYN
<i>Gastric acid secretion</i>	0.039	CAMK2D;GNAI3;ATP1A3;ITPR3;CALML3;GAST;SLC9A1
<i>Viral myocarditis</i>	0.040	FYN;HLA-A;HLA-DRB3;ITGAL;HLA-G;HLA-DRB1
<i>Pertussis</i>	0.042	MAPK11;IRAK1;SFTPA2;CFL1;GNAI3;CALML3;C4BPB

*continua*

**Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41** continuação

Turquoise	TNF signaling pathway	0.000	EDN1;CCL20;LIF;PIK3R1;FOS;CXCL3;TNF;CXCL2;ICAM1;TNFRSF1A;NFKBIA;ITCH;CASP3;BCL3;MAP3K8;FADD;DNM1L;JUNB;BIRC3
	NF-kappa B signaling pathway	0.002	NFKBIA;GADD45B;PLAU;TRIM25;LTB;TICAM1;CXCL2;TNF;TNFRSF1A;RELB;BIRC3;ICAM1
	Metabolic pathways	0.002	PYGB;NDUFA11;ADPGK;COMT;NDST2;ZNRD1;TDO2;HYAL1;LIPG;ME3;UPP1;PGM1;UGT2B10;ACAD8;PDHX;DGAT1;GPT2;DNMT3A;ALG2;ALG10;ATP5F1;MUT;PRODH2;ITPKC;NDUFS8;PHOSPHO2;PIGA;IMPAD1;NDUFS4;UGT2B4;PIK3C3;B3GNT1;PLCB1;ALDH7A1;PAFAH1B1;BCAT2;PDXK;MAOB;RPE;PRIM1;NDUFB2;DLST;ATP5O;FUT4;ACAT1;LDHB;SGPL1;CYP2B6;MAT2A;RDH10;INPP5E;UAP1L1;GCNT2;POLR2F;PCK1;POLR2J;ARSB;ATP6V0A1;DUT;FAHD1;GK;GCH1;POLR3GL;AGL;EPHX2;HSD3B1;MTHFR;DHCR24;CYP8B1;PRDX6;COQ6;DAK;DHFR;SQLE;COQ3;UCK2;CPS1;IMPDH1;POLE2;CYP1A1;PPT1;LPIN3
	Ubiquitin mediated proteolysis	0.003	PIAS3;UBE2H;CUL7;UBE3C;UBE2G1;RNF7;BRCA1;PIAS2;RHOBTB2;CUL4A;ITCH;SOCS1;CDC34;UBE2R2;BIRC3
	Cell cycle	0.003	CDKN1C;CDKN1A;PCNA;GADD45B;MCM7;CDKN2A;TTK;PKMYT1;CDC25A;GADD45G;WEE1;CHEK1;E2F1;BUB3
	Apoptosis	0.003	GADD45B;PIK3R1;FOS;TNF;EIF2S1;GADD45G;TNFRSF1A;BBC3;NFKBIA;CASP3;FADD;BAK1;BID;BIRC3;CTSB
	Bladder cancer	0.003	CDKN1A;RASSF1;CDKN2A;MMMP1;E2F1;BRAF;HBEGF
	Protein export	0.004	IMMP1L;IMMP2L;SPCS1;SRP72;SRP54
	Legionellosis	0.005	NFKBIA;CLK1;SAR1A;CASP3;CXCL3;CXCL2;TNF;HSPA1A

continua

**Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41** continuação

<i>p53 signaling pathway</i>	0.006	CDKN1A;TP53I3;GADD45B;CDKN2A;CASP3;CHEK1;BID;BBC3;GADD45G
<i>Pathways in cancer</i>	0.006	CDKN1A;PTGER2;LAMA4;PIK3R1;ADCY6;HSP90B1;RASSF1;RXRA;CASP3;GNA12;E2F1;FADD;BID;EGLN2;CDKN2A;TPM3;MP1;F2R;BRAF;LAMB1;FOS;NG12;MLH1;PIAS2;MSH6;NFKBIA;BMP2;PAX8;TRAF4;PLCB1;BIRC3
<i>Peroxisome</i>	0.007	PEX16;SLC25A17;EPHX2;MPV17L;PXMP4;PEX10;MPV17;SOD2;PEX13;CRAT
<i>Nucleotide excision repair</i>	0.007	RFC5;CUL4A;PCNA;POLE2;GTF2H1;ERCC5;MNAT1
<i>Hepatitis C</i>	0.012	CDKN1A;BRAF;PPP2R2A;PIK3R1;TICAM1;TNF;EIF2S1;TNFRSF1A;NFKBIA;RXRA;PPP2R1B;CLDN23;LDLR
<i>mRNA surveillance pathway</i>	0.012	CPSF4;PABPN1;PPP2R1B;CSTF3;CSTF1;SYMPK;CSTF2T;PPP2R5D;PPP2R2A;PPP2R3A
<i>Complement and coagulation cascades</i>	0.014	VWF;PLAU;F12;FGG;F2R;TFPI;CFB;C2;MBL2
<i>Ubiquinone and other terpenoid-quinone biosynthesis</i>	0.014	VKORC1;COQ3;COQ6
<i>Non-alcoholic fatty liver disease (NAFLD)</i>	0.015	MLXIP;NDUFA11;NDUFB2;PIK3R1;TNF;EIF2S1;TNFRSF1A;MLXIPL;ITCH;RXRA;NDUFS8;CASP3;NDUFS4;BID
<i>Alzheimer's disease</i>	0.016	NDUFA11;NDUFB2;ATP5F1;ATP5O;TNF;TNFRSF1A;BACE1;NDUFS8;CASP3;NDUFS4;FADD;PLCB1;BID;ATF6;CDK5R1
<i>Herpes simplex infection</i>	0.018	MED8;TAP1;FOS;TICAM1;TNF;EIF2S1;TNFRSF1A;NFKBIA;PER2;CDC34;UBE2R2;CASP3;GTF2IRD1;FADD;PVRL2;GTF2I
<i>Vasopressin-regulated water reabsorption</i>	0.018	RAB5C;DYNLL1;DYNLL2;AQP3;ADCY6;DYNC1I1
<i>Basal transcription factors</i>	0.020	GTF2A2;TAF11;GTF2H1;GTF2IRD1;MNAT1;GTF2I
<i>Small cell lung cancer</i>	0.023	NFKBIA;RXRA;TRAF4;LAMA4;E2F1;LAMB1;PIK3R1;PIAS2;BIRC3
<i>Mismatch repair</i>	0.023	RFC5;MSH6;PCNA;MLH1

continua



**Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41** continuação

	<i>MAPK signaling pathway</i>	0.023	<i>DUSP5;DUSP2;JUND;DUSP3;GADD45B;DUSP1;BRAF;FOS;GN G12;TNF;DUSP6;GADD45G;RE LB;TNFRSF1A;RAP1A;TAOK3;C ASP3;GNA12;MAP3K8;HSPA1A</i>
	<i>AGE-RAGE signaling pathway in diabetic complications</i>	0.024	<i>EGR1;DIAPH1;EDN1;CASP3;PI M1;PIK3R1;PLCB1;TNF;AGER;I CAM1</i>
	<i>mTOR signaling pathway</i>	0.025	<i>STK11;RRAGC;RPS6;AKT1S1;B RAF;PIK3R1;TNF</i>
	<i>Folate biosynthesis</i>	0.027	<i>DHFR;GCH1;GGH</i>
	<i>Toxoplasmosis</i>	0.028	<i>NFKBIA;SOCS1;CASP3;LAMA4; LAMB1;PIK3R1;TNF;LDLR;HSP A1A;BIRC3;TNFRSF1A</i>
	<i>DNA replication</i>	0.028	<i>RFC5;PCNA;MCM7;POLE2;PRI M1</i>
	<i>Chagas disease (American trypanosomiasis)</i>	0.029	<i>NFKBIA;PPP2R1B;PPP2R2A;PI K3R1;FADD;FOS;TICAM1;PLCB 1;TNF;TNFRSF1A</i>
	<i>Sphingolipid signaling pathway</i>	0.031	<i>SGPL1;PPP2R1B;GNA12;PPP2 R5D;PPP2R2A;PIK3R1;PPP2R3 A;PLCB1;BID;TNF;TNFRSF1A</i>
	<i>Vitamin B6 metabolism</i>	0.031	<i>PDXK;PHOSPHO2</i>
	<i>Lysosome</i>	0.036	<i>CD164;CLN3;SORT1;HYAL1;M6 PR;PPT1;AP3S2;CD68;ARSB;A TP6V0A1;CTSB</i>
	<i>Dorso-ventral axis formation</i>	0.039	<i>PIWIL2;NOTCH2;SPIRE2;CPEB 4</i>
	<i>Tryptophan metabolism</i>	0.042	<i>MAOB;TDO2;CYP1A1;ALDH7A1 ;ACAT1</i>
	<i>Pyruvate metabolism</i>	0.042	<i>LDHB;ME3;PCK1;ALDH7A1;AC AT1</i>
Pink	<i>Protein processing in endoplasmic reticulum</i>	0.002	<i>NSFL1C;UBE2D4;PRKCSH;BA G1;MAN1C1;DERL2;PDIA4</i>
	<i>Pathogenic Escherichia coli infection</i>	0.002	<i>ROCK2;TUBB;ARPC5;YWHAZ</i>
	<i>Porphyrin and chlorophyll metabolism</i>	0.008	<i>HMBS;GUSB;UGT2B7</i>
	<i>Drug metabolism - other enzymes</i>	0.011	<i>GUSB;ITPA;UGT2B7</i>
	<i>Huntington's disease</i>	0.045	<i>NDUFB8;GPX1;POLR2C;VDAC 2;COX6B1</i>

continua

**Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41** continuação

<i>Brown</i>	<i>Systemic lupus erythematosus</i>	0.000	<i>HIST1H2BJ;HIST1H2BL;HIST1H2BK;HIST2H2AC;SNRPD1;HIST1H2AE;HIST1H2AD;HIST1H3H;HIST3H2BB;HIST1H3B;HIST1H3C;HIST1H2AC;HIST1H3D;HIST1H2AB;HIST1H2AM;HIST1H2AL;HIST1H4K;H3F3B;HIST1H4L;HIST1H2AH;HIST1H2AK;HIST1H2AJ;H2AFX;ACTN4;HIST1H4A;HIST2H3A;HIST1H4H;HIST1H2BE;HIST1H4I;HIST1H2BH;HIST1H2BB;HIST1H4D;HIST1H2BD;HIST1H2BC</i>
	<i>Alcoholism</i>	0.000	<i>HIST1H2BJ;HIST1H2BL;HIST1H2BK;HIST2H2AC;GNAI2;NRAS;GNG7;HIST1H2AE;HIST1H2AD;HIST1H3H;HIST3H2BB;HIST1H3B;HIST1H3C;HIST1H2AC;HIST1H3D;HIST1H2AB;HIST1H2AM;HIST1H2AL;HIST1H4K;H3F3B;HIST1H4L;HIST1H2AH;HIST1H2AK;HIST1H2AJ;H2AFX;GRIN2C;GNG13;HIST1H4A;HIST2H3A;HIST1H4H;HIST1H2BE;HIST1H4I;HIST1H2BH;HIST1H2BB;HIST1H4D;HIST1H2BD;HIST1H2BC</i>
	<i>Viral carcinogenesis</i>	0.000	<i>YWHAE;HIST1H2BJ;SRC;HIST1H2BL;HIST1H2BK;GTF2E1;CDC42;NRAS;HIST3H2BB;YWHAG;STAT5B;HIST1H4K;HIST1H4L;STAT3;SCRIB;ACTN4;DDB1;HIST1H4A;MDM2;IRF7;HIST1H4H;HIST1H2BE;HIST1H4I;HIST1H2BH;HIST1H2BB;HIST1H4D;HIST1H2BD;HIST1H2BC</i>
	<i>Ubiquitin mediated proteolysis</i>	0.000	<i>UBE2B;CUL5;UBE2C;UBE2D2;UBE2E3;SIAH1;UBE4A;UBE2E2;CBLB;RBX1;UBE2J1;DDB1;UBE2Q1;CDC27;MDM2;BIRC6;TRIM37;STUB1;ANAPC1</i>
	<i>Protein processing in endoplasmic reticulum</i>	0.000	<i>UFD1L;BCAP31;TRAM1;UBE2D2;SSR2;UBE2E3;DNAJB12;DERL1;UBE2E2;RAD23A;RBX1;UBE2J1;CANX;CAPN2;PLAA;SSR1;STUB1;SEC24C;MBTPS2</i>

continua

**Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41** continuação

<i>Proteoglycans in cancer</i>	0.000	MAP2K2;ROCK1;CAV2;SRC;RDX;STAT3;FRS2;CBLB;IQGAP1;MAPK14;MMP9;ACTB;CDC42;NRAS;SMO;PDCD4;MDM2;CTNNB1;FLNA
<i>Metabolic pathways</i>	0.000	PIGS;ALAS1;PIGO;HIBADH;NDUFA12;COX7C;GCSH;GMPPB;LIPC;HYAL3;PSPH;ST6GAL1;TPI1;ACSL1;PGAM1;GAA;DCXR;UGT2B15;BCKDHB;ALG3;CTPS2;CEL;MINPP1;MTAP;NME6;NME7;ATP6V1B2;UQCRC2;PIGL;DLAD;GAPDH;GNE;TSTA3;LIPT1;DNMT1;ISYNA1;NDUFB7;ADH1C;HSD17B12;HSD17B6;MAT2B;AGPAT1;LIAS;POLD4;PMVK;TWISTNB;RDH16;GCNT3;POLR2G;LTA4H;FUK;ATP6V1D;PDHA1;PTGES2;PMM2;B3GALT6;MTHFS;UCKL1;ALDH4A1;RPIA;GCLC;COQ2;SUCLA2;QARS;POLR3B;POLR3H;ACO2;ALDH18A1;LDHAL6A
<i>Bacterial invasion of epithelial cells</i>	0.002	CDC42;ARHGAP10;CAV2;SRC;CLTB;ILK;CTNNB1;CBLB;ACTB
<i>Glycolysis / Gluconeogenesis</i>	0.003	MINPP1;TPI1;PDHA1;ADH1C;PGAM1;DLAD;GAPDH;LDHAL6A
<i>Pyruvate metabolism</i>	0.003	PDHA1;GLO1;ACYP1;HAGH;DLAD;LDHAL6A
<i>Homologous recombination</i>	0.004	TOP3B;POLD4;RPA3;TOP3A;RPA1
<i>Regulation of actin cytoskeleton</i>	0.006	APC2;MAP2K2;ROCK1;SRC;ITGB4;RDX;ACTN4;IQGAP1;IQGAP2;SSH3;ACTB;CDC42;NRAS;ARRHGEF4;PAK6;ITGB8
<i>Pyrimidine metabolism</i>	0.006	POLD4;POLR3B;NME6;NME7;ENTPD5;TWISTNB;CTPS2;POLR3H;POLR2G;UCKL1
<i>Adherens junction</i>	0.006	CDC42;SRC;CTNNB1;ACTN4;IQGAP1;NLK;ACTB;TGFB2
<i>Biosynthesis of amino acids</i>	0.006	RPIA;TPI1;PGAM1;ALDH18A1;ACO2;MAT2B;GAPDH;PSPH
<i>Fructose and mannose metabolism</i>	0.006	GMPPB;TSTA3;TPI1;PMM2;FUK
<i>HTLV-I infection</i>	0.008	APC2;STAT5B;MAP2K4;CRTC2;RANBP3;CRTC3;TGFB2;ELK4;CCNB2;POLD4;NRAS;PTTG1;CDC27;CANX;TP53INP1;CTNNB1;SLC25A4;ANAPC1

continua

**Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41** continuação

<i>Nucleotide excision repair</i>	0.007	<i>DDB1;POLD4;RPA3;RPA1;RAD23A;RBX1</i>
<i>Focal adhesion</i>	0.008	<i>VASP;ROCK1;CAV2;SRC;ITGB4;ILK;ACTN4;ACTB;CDC42;ZYX;CAPN2;CTNNB1;FLNA;PAK6;ITGB8</i>
<i>Ribosome</i>	0.013	<i>RPS14;MRPS9;MRPL19;RPL13;MRPL14;RPL23A;RPL9;RPL39;RPS10;MRPS18C;MRPL11</i>
<i>Leukocyte transendothelial migration</i>	0.012	<i>VASP;CDC42;ROCK1;CLDN7;CTNNB1;ACTN4;MAPK14;MMP9;ACTB;GNAI2</i>
<i>Fanconi anemia pathway</i>	0.013	<i>TOP3B;RPA3;TOP3A;RPA1;HESS1;POLK</i>
<i>Tight junction</i>	0.014	<i>CDC42;NRAS;SRC;CLDN7;CTNNB1;ACTN4;ASH1L;PRKCZ;TJAP1;ACTB;GNAI2</i>
<i>Glycine, serine and threonine metabolism</i>	0.015	<i>GCSH;ALAS1;PGAM1;DLD;PSPH</i>
<i>Cell cycle</i>	0.017	<i>YWHAE;CCNB2;PTTG1;CDC27;MDM2;SMC3;YWHAG;ANAPC1;RBX1;CDC14B</i>
<i>Bladder cancer</i>	0.017	<i>NRAS;MAP2K2;SRC;MDM2;MMP9</i>
<i>Carbon metabolism</i>	0.024	<i>RPIA;TPI1;PDHA1;SUCLA2;PGAM1;ACO2;DLD;GAPDH;PSPH</i>
<i>VEGF signaling pathway</i>	0.025	<i>CDC42;NRAS;MAP2K2;MAPKAPK3;SRC;MAPK14</i>
<i>Proteasome</i>	0.023	<i>PSMB5;PSMD4;PSMB3;PSME4;PSMB10</i>
<i>Cysteine and methionine metabolism</i>	0.025	<i>MTAP;GCLC;DNMT1;MAT2B;LDHAL6A</i>
<i>Citrate cycle (TCA cycle)</i>	0.024	<i>PDHA1;SUCLA2;ACO2;DLD</i>
<i>Endocytosis</i>	0.032	<i>CAV2;SRC;ZFYVE20;CLTB;CBLB;STAM;SNF8;PRKCZ;TGFB2;CDC42;CAPZA1;MDM2;CHMP6;FOLR2;SMAP1;CHMP5</i>
<i>Shigellosis</i>	0.033	<i>CDC42;ROCK1;SRC;UBE2D2;MAPK14;ACTB</i>
<i>Amino sugar and nucleotide sugar metabolism</i>	0.032	<i>GMPPB;TSTA3;PMM2;FUK;GNE</i>
<i>Propanoate metabolism</i>	0.029	<i>SUCLA2;BCKDHB;DLD;LDHAL6A</i>
<i>RNA polymerase</i>	0.029	<i>POLR3B;TWISTNB;POLR3H;POLR2G</i>
<i>Base excision repair</i>	0.032	<i>POLD4;OGG1;APEX2;UNG</i>
<i>Oocyte meiosis</i>	0.039	<i>YWHAE;CCNB2;PTTG1;CDC27;PPP2R5B;SMC3;YWHAG;ANAPC1;RBX1</i>

continua

**Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41** continuação

	<i>Epithelial cell signaling in Helicobacter pylori infection</i>	0.040	CDC42;MAP2K4;SRC;ATP6V1B2;MAPK14;ATP6V1D
	<i>ErbB signaling pathway</i>	0.042	STAT5B;MAP2K4;NRAS;MAP2K2;SRC;PAK6;CBLB
	<i>Endometrial cancer</i>	0.043	APC2;NRAS;MAP2K2;ILK;CTNNB1
	<i>Hepatitis B</i>	0.045	DDB1;STAT5B;MAP2K4;NRAS;MAP2K2;SRC;STAT3;IRF7;MMP9;IFNAR1
Yellow	<i>RNA transport</i>	0.000	CYFIP1;EIF5B;POP7;RBM8A;PO1;THOC3;CASC3;EEF1A1;NUP85;RPP21;EIF4EBP1;RNPS1;GEMIN7;EIF2B1;EIF4E2;EIF1B;EIF4B;KPNB1
	<i>Huntington's disease</i>	0.000	DNAH11;NDUFA6;UQCRB;HDAC1;NDUFA3;NDUFA1;COX5A;CASP9;POLR2B;NDUFS5;NDUFA1;PPIF;EP300;UQCRFS1;POLR2H
	<i>Epstein-Barr virus infection</i>	0.000	RB1;MAP2K3;HSPA8;GSK3B;HDAC1;POLR3A;PSMD7;POLR3C;POLR2B;POLR3D;AKT2;CSNK2B;EP300;RIPK1;POLR2H
	<i>Non-homologous end-joining</i>	0.000	FEN1;RAD50;XRCC4;XRCC5
	<i>Non-alcoholic fatty liver disease (NAFLD)</i>	0.001	SREBF1;GSK3B;XBP1;NDUFA6;UQCRB;AKT2;NDUFS5;NDUFA3;NDUFAB1;NDUFA1;UQCRFS1;COX5A
	<i>Oxidative phosphorylation</i>	0.001	NDUFA6;UQCRB;NDUFS5;NDUFA3;NDUFAB1;NDUFA1;ATP6V1H;UQCRFS1;ATP6V1E1;COX5A;ATP5J2
	<i>Spliceosome</i>	0.001	HSPA8;LSM6;RBM8A;BUD31;PCBP1;CDC5L;THOC3;WBP11;CRNKL1;RBMX;CTNNBL1
	<i>Parkinson's disease</i>	0.001	CASP9;NDUFA6;UQCRB;NDUFS5;NDUFA3;NDUFAB1;PPIF;NDUFA1;HTRA2;UQCRFS1;COX5A
	<i>RNA polymerase</i>	0.001	POLR3A;POLR3C;POLR2B;POLR3D;POLR2H

continua

**Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41** continuação

<i>Metabolic pathways</i>	0.002	<i>B4GALT3;MOCS2;EPRS;NADSYN1;GBA2;SGSH;IMPA1;ALDH2;GUK1;PPAT;HYAL2;UQCRFS1;PGLS;ATP6V1E1;ACAA1;SEPHS1;NME1;ATP5J2;PIGC;HADHB;NDUFS5;GART;MTMR1;UQCRB;FECH;SHMT1;COX5A;HSD17B8;TRAK2;POLD3;UGT8;PPCS;POLR2B;ATP6V1H;POLR2H;MPST;NDUFA6;NDUFA3;IDH1;NDUFA1;TRIT1;AGMAT;CS;POLR3A;POLR3C;POLR3D;NDUFAB1;DHCR7;PFKM</i>
<i>Prostate cancer</i>	0.002	<i>RB1;CASP9;GSK3B;CCNE1;AKT2;EP300;RAF1;FGFR1</i>
<i>DNA replication</i>	0.002	<i>POLD3;FEN1;RNASEH1;RPA2;MCM4</i>
<i>Legionellosis</i>	0.003	<i>EEF1A1;CASP9;HSPA8;BCL2L13;SAR1B;MYD88</i>
<i>Ribosome</i>	0.003	<i>RPL21;RPL10;RPL22;MRPS18A;RPL26;RPL29;RPS27A;RPL17;MRPS5;RPL7</i>
<i>Apoptosis</i>	0.003	<i>CASP9;DFFA;CASP6;CASP10;AKT2;BIRC5;HTRA2;RIPK1;RAF1;PTPN13</i>
<i>Influenza A</i>	0.006	<i>MAP2K3;DNAJC3;CASP9;HSPA8;GSK3B;OAS1;AKT2;EP300;RAF1;PML;MYD88</i>
<i>Base excision repair</i>	0.010	<i>POLD3;MBD4;FEN1;APEX1</i>
<i>Thyroid hormone signaling pathway</i>	0.011	<i>CASP9;GSK3B;HDAC1;AKT2;ATP1A4;DIO3;EP300;RAF1</i>
<i>Alzheimer's disease</i>	0.012	<i>CASP9;GSK3B;NDUFA6;UQCRB;NDUFS5;NDUFA3;NDUFAB1;NDUFA1;UQCRFS1;COX5A</i>
<i>Adherens junction</i>	0.012	<i>CSNK2B;EP300;PVRL3;ACP1;VCL;FGFR1</i>
<i>Hepatitis B</i>	0.014	<i>RB1;CASP9;CASP10;CCNE1;AKT2;BIRC5;EP300;RAF1;MYD88</i>
<i>Cell cycle</i>	0.015	<i>RB1;GSK3B;ANAPC7;CCNE1;HDAC1;MCM4;EP300;SFN</i>
<i>Cardiac muscle contraction</i>	0.015	<i>UQCRB;TNNC1;TPM1;ATP1A4;UQCRFS1;COX5A</i>
<i>Purine metabolism</i>	0.016	<i>POLD3;POLR3A;POLR3C;GUK1;POLR2B;POLR3D;PPAT;POLR2H;GART;NME1</i>
<i>Pyrimidine metabolism</i>	0.019	<i>POLD3;POLR3A;POLR3C;POLR2B;POLR3D;POLR2H;NME1</i>
<i>mTOR signaling pathway</i>	0.019	<i>RRAGA;AKT2;EIF4EBP1;EIF4E2;EIF4B</i>
<i>Mismatch repair</i>	0.021	<i>POLD3;EXO1;RPA2</i>

continua

**Anexo 3 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/EH41** conclusão

<i>Colorectal cancer</i>	0.022	<i>CASP9;GSK3B;AKT2;BIRC5;RAF1</i>
<i>Hepatitis C</i>	0.022	<i>GSK3B;SCARB1;OCLN;CLDN3;OAS1;AKT2;RIPK1;RAF1</i>
<i>Cytosolic DNA-sensing pathway</i>	0.025	<i>POLR3A;POLR3C;POLR3D;RIPK1;POLR2H</i>
<i>Sulfur relay system</i>	0.026	<i>MPST;MOCS2</i>
<i>Fatty acid degradation</i>	0.026	<i>HADHB;CPT1A;ALDH2;ACAA1</i>
<i>Aminoacyl-tRNA biosynthesis</i>	0.028	<i>DARS;KARS;TARS;GARS;EPRS</i>
<i>Insulin signaling pathway</i>	0.028	<i>SREBF1;GSK3B;EXOC7;AKT2;EIF4EBP1;PHKB;RAF1;EIF4E2</i>
<i>Central carbon metabolism in cancer</i>	0.029	<i>AKT2;IDH1;RAF1;PFKM;FGFR1</i>
<i>Homologous recombination</i>	0.038	<i>POLD3;RAD50;RPA2</i>
<i>Endometrial cancer</i>	0.045	<i>CASP9;GSK3B;AKT2;RAF1</i>
<i>Glutathione metabolism</i>	0.045	<i>GSTA5;GSTA4;GSTA2;IDH1</i>

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472**

Módulo	KEGG	valor-p	Genes
Yellow	Cell cycle	0.000	CDKN2D;ANAPC13;SMAD3;CDKN2A;BUB1B;TTK;CDC25C;PKMYT1;CDC25A;CDC20;CCNA2;CCNB2;PTTG1;PTTG2;CHEK1;E2F1;ABL1;MCM3;E2F2;E2F3;MCM5;BUB1;MAD2L1;MCM2
	HTLV-I infection	0.000	MAP3K3;SMAD3;FZD4;CDKN2A;ADCY3;BUB1B;CDC20;CCNB2;PTTG1;PTTG2;CHEK1;CANX;E2F1;E2F2;E2F3;TCF3;ATF3;WNT3;MAD2L1
	Fanconi anemia pathway	0.000	APITD1;RAD51;RMI1;UBE2T;USP1;FANCA;BRCA1;FANCE
	Oocyte meiosis	0.000	CDC20;ANAPC13;CCNB2;PTTG1;PTTG2;ADCY3;ITPR3;CDC25C;PKMYT1;BUB1;MAD2L1;AURKA
	Ubiquitin mediated proteolysis	0.001	CDC20;ANAPC13;PPIL2;UBE2Q1;UBE2C;UBE2S;MGRN1;SYVN1;BRCA1;TRIM32;RHOTB2
	Progesterone-mediated oocyte maturation	0.001	CCNA2;ANAPC13;CCNB2;ADCY3;CDC25C;PKMYT1;BUB1;CDC25A;MAD2L1
	p53 signaling pathway	0.001	CCNB2;RRM2;CDKN2A;CCNG2;CCNG1;CHEK1;BID
	DNA replication	0.002	FEN1;LIG1;MCM3;MCM5;MCM2
	Vitamin digestion and absorption	0.002	SLC5A6;SCARB1;APOA1;PLB1
	MicroRNAs in cancer	0.003	NOTCH2;CDKN2A;CDCA5;PRKCA;KIF23;BRCA1;CDC25C;CDC25A;STMN1;CCNG1;E2F1;ABL1;E2F2;E2F3;WNT3;EZH2
	Bladder cancer	0.003	CDKN2A;MMP1;E2F1;E2F2;E2F3
	Homologous recombination	0.005	RAD51;RAD54L;NBN;RAD54B
	Pancreatic cancer	0.005	SMAD3;RAD51;CDKN2A;E2F1;E2F2;E2F3
	2-Oxocarboxylic acid metabolism	0.007	GOT1;GPT2;BCAT2
	Chronic myeloid leukemia	0.009	SMAD3;CDKN2A;E2F1;ABL1;E2F2;E2F3
	African trypanosomiasis	0.010	LAMA4;IL18;APOA1;PRKCA
	Non-small cell lung cancer	0.011	CDKN2A;E2F1;E2F2;E2F3;PRKCA
	Glioma	0.021	CDKN2A;E2F1;E2F2;E2F3;PRKCA
	Amino sugar and nucleotide sugar metabolism	0.029	MPI;GMPPA;UAP1;GNE

continua



**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

	<i>Notch signaling pathway</i>	0.029	NCOR2;NOTCH2;MAML1;RFNG
	<i>Thyroid hormone synthesis</i>	0.029	CANX;ADCY3;ITPR3;PRKCA;TTF2
	<i>Hippo signaling pathway</i>	0.033	SMAD3;FZD4;WWC1;ID1;BIRC5;NF2;AFP;WNT3
	<i>Pathways in cancer</i>	0.034	CSF1R;SMAD3;FZD4;CDKN2A;MMP1;LAMA4;ADCY3;PRKCA;RAD51;ABL1;E2F1;BIRC5;E2F2;E2F3;BID;WNT3
	<i>Glutathione metabolism</i>	0.037	GSTM3;RRM1;RRM2;GSS
	<i>Fatty acid biosynthesis</i>	0.038	ACSL4;ACACA
	<i>Antigen processing and presentation</i>	0.039	KIR2DS2;NFYC;RFX5;CANX;PSME2
<i>Blue</i>	<i>Spliceosome</i>	0.000	SF3B2;DDX5;RBM25;NCBP2;WBP11;EFTUD2;LSM7;PHF5A;NHP2L1;SYF2;SNRPG;SNRPE;RP9;SNRPB2;SNRPD3;TXNL4A;RBMX;SMNDC1
	<i>Alzheimer's disease</i>	0.000	NDUFA9;NDUFB9;NDUFA6;NDUFA3;ATP5C1;GRIN2C;ATP5F1;COX5A;TNFRSF1A;PPP3CA;COX7A2L;PLCB3;NDUFS5;CASP3;NDUFS1;UQCRC2;FADD;GAPDH;SNCA
	<i>Huntington's disease</i>	0.000	NDUFA9;NDUFB9;GPX1;NDUFA6;NDUFA3;ATP5C1;CLTA;AP2B1;ATP5F1;COX5A;COX7A2L;PLCB3;POLR2B;NDUFS5;CASP3;TFAM;VDAC2;NDUFS1;UQCRC2;POLR2L
	<i>Systemic lupus erythematosus</i>	0.000	HIST1H2AM;HIST1H4K;HIST1H2AH;HIST1H2AK;HIST1H2BL;HIST1H4A;HIST2H3A;HIST1H2AE;HIST1H2AD;HIST1H4J;SNRPD3;HLA-DRB1;HLA-DQA1;HIST1H2BD;HIST1H2AB
	<i>Non-alcoholic fatty liver disease (NAFLD)</i>	0.000	NDUFA9;NDUFB9;NDUFA6;NDUFA3;PRKAG2;PRKAB1;COX5A;TNFRSF1A;CDC42;COX7A2L;RXRA;PIK3CA;NDUFS5;CASP3;NDUFS1;UQCRC2
	<i>Alcoholism</i>	0.000	HIST1H2AM;ATF2;HDAC5;HIST1H4K;HIST1H2AH;HIST1H2AK;HIST1H2BL;GRIN2C;CAMKK2;GNAO1;HIST1H4A;HIST2H3A;HIST1H2AE;HIST1H2AD;HIST1H4J;HIST1H2BD;HIST1H2AB

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>Parkinson's disease</i>	0.000	NDUFA9;NDUFB9;NDUFA6;NDUFA3;ATP5C1;ATP5F1;COX5A;COX7A2L;NDUFS5;CASP3;VDAC2;NDUFS1;UQCRC2;SNCA
<i>MAPK signaling pathway</i>	0.001	ATF2;DUSP3;DUSP1;PLA2G4C;PDGFA;ARRB1;STK4;MAPK8IP3;ELK1;CACNA1G;TNFRSF1A;RAP1B;FGF5;CDC42;PPP3CA;DUSP10;MAPKAPK3;CASP3;MAP3K13
<i>Viral carcinogenesis</i>	0.001	ATF2;HDAC5;GTF2A2;HIST1H4K;GTF2B;HIST1H2BL;HLA-B;SCRIB;CDC42;HIST1H4A;RBL1;PIK3CA;CASP3;HIST1H4J;IL6ST;HIST1H2BD
<i>Oxidative phosphorylation</i>	0.001	NDUFA9;NDUFB9;COX7A2L;NDUFA6;ATP6AP1;NDUFS5;NDUFA3;ATP5C1;NDUFS1;UQCRC2;ATP5F1;COX5A
<i>Chagas disease (American trypanosomiasis)</i>	0.002	GNAO1;PPP2CA;PLCB3;PIK3CA;CL3L3;PPP2R2D;CFLAR;FADD;IL2;TNFRSF1A
<i>Calcium signaling pathway</i>	0.002	CAMK2D;SPHK2;TNNC1;F2R;GRIN2C;CACNA1G;PPP3CA;PLCB3;P2RX3;P2RX2;ITPKA;VDAC2;PTK2B;CAMK2G
<i>Metabolic pathways</i>	0.004	ACAA2;CMAS;PIGO;ATP5C1;CNDP2;LIPC;ALDH2;PIP5KL1;PDHX;ATP6AP1;SPHK2;PLA2G4C;TALDO1;NME3;ATP5F1;CHIT1;HADHA;MTAP;PLCB3;NME7;NDUFS5;HMBS;ITPKA;NDUFS1;UQCRC2;PIGL;GAPDH;ACOT4;ALDH9A1;PON3;NDUFB9;MTMR2;FECH;COX5A;AGPAT3;ADH5;LDHB;GANAB;INPP5A;RDH11;CYP11A1;POLR2B;RDH10;AFMID;POLR2L;NDUFA9;NDUFA6;GALNT3;TPK1;NDUFA3;CYP51A1;CRLS1;AMACR;PSAT1;PFKP;LPIN3;CDS2
<i>RNA transport</i>	0.004	EIF5B;EIF2B3;NCBP2;NMD3;EIF2S2;FXR2;XPO1;SUMO1;NUP54;EIF1B;EIF4G2;WIBG;PAIP1
<i>Endocytosis</i>	0.005	RAB4A;F2R;HLA-B;CLTA;VPS26A;ARRB1;VPS26B;AP2B1;PML;ARFGAP1;SNX3;CDC42;RAB11FIP1;SNX2;PIP5KL1;CHMP7;RAB8A
<i>Glucagon signaling pathway</i>	0.006	ATF2;PPP3CA;LDHB;PLCB3;CAMK2D;PRMT1;PRKAG2;PRKAB1;CAMK2G

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>mRNA surveillance pathway</i>	0.010	PPP2CA;SMG1;NCBP2;CPSF3;PP2R2D;CSTF2;WIBG;SMG6
<i>Graft-versus-host disease</i>	0.010	HLA-B;KIR3DL2;IL2;HLA-DRB1;HLA-DQA1
<i>HTLV-I infection</i>	0.011	ATF2;IL15RA;HLA-B;PDGFA;ELK1;IL2;ANAPC11;TNFRSF1A;PPP3CA;XPO1;PIK3CA;LCK;TSPO;VDAC2;HLA-DRB1;HLA-DQA1
<i>Viral myocarditis</i>	0.012	CASP3;HLA-B;ABL2;HLA-DQA1;EIF4G2;HLA-DRB1
<i>Natural killer cell mediated cytotoxicity</i>	0.013	NCR2;PPP3CA;TYROBP;PIK3CA;LCK;CASP3;HLA-B;PTK2B;TNFRSF10A;KIR3DL2
<i>Fatty acid degradation</i>	0.014	HADHA;ACAA2;ALDH2;ADH5;ALDH9A1
<i>Thyroid cancer</i>	0.014	TCF7L2;TCF7L1;RXRA;TFG
<i>Tuberculosis</i>	0.014	PPP3CA;CAMK2D;CLEC4M;LAMP1;ATP6AP1;SPHK2;CASP3;FADD;CAMK2G;HLA-DRB1;HLA-DQA1;TNFRSF1A
<i>VEGF signaling pathway</i>	0.014	CDC42;PPP3CA;PIK3CA;MAPKAPK3;SPHK2;PLA2G4C
<i>Ribosome</i>	0.015	RPL5;RPL21;MRPS12;RPL23A;RPL26;RPL15;RPL10A;MRPL32;RPS12;MRPL22
<i>Oxytocin signaling pathway</i>	0.015	GNAO1;PPP3CA;PLCB3;CAMK2D;PIK3CA;PLA2G4C;PRKAG2;PRKAB1;ELK1;CAMK2G;CAMKK2
<i>beta-Alanine metabolism</i>	0.018	HADHA;ALDH2;ALDH9A1;CNBP2
<i>AMPK signaling pathway</i>	0.021	PPP2CA;STK11;PIK3CA;PPP2R2D;PRKAG2;PRKAB1;PFKP;RAB8A;CAMKK2
<i>Long-term potentiation</i>	0.021	RAP1B;PPP3CA;PLCB3;CAMK2D;GRIN2C;CAMK2G
<i>Glycolysis / Gluconeogenesis</i>	0.022	LDHB;ALDH2;GAPDH;PFKP;ADH5;ALDH9A1
<i>Regulation of actin cytoskeleton</i>	0.024	ITGB1;NCKAP1;ITGA3;F2R;PDGFA;ITGAE;IQGAP2;FGF5;CDC42;GNA13;PIK3CA;BCAR1;PAK4
<i>Amyotrophic lateral sclerosis (ALS)</i>	0.025	PPP3CA;GPX1;CASP3;GRIN2C;TNFRSF1A
<i>Dopaminergic synapse</i>	0.026	GNAO1;PPP2CA;ATF2;PPP3CA;PLCB3;CAMK2D;PPP2R2D;CAMK2G;DRD4
<i>Ribosome biogenesis in eukaryotes</i>	0.027	XPO1;NHP2L1;NVL;CSNK2A2;HEATR1;WDR75;NMD3
<i>Adipocytokine signaling pathway</i>	0.027	STK11;RXRA;PRKAG2;PRKAB1;TNFRSF1A;CAMKK2

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

	<i>Inositol phosphate metabolism</i>	0.028	<i>MTMR2;PLCB3;INPP5A;PIK3CA;PIP5KL1;ITPKA</i>
	<i>Autoimmune thyroid disease</i>	0.029	<i>HLA-B;IL2;HLA-DQA1;HLA-DRB1;TSHB</i>
	<i>GnRH signaling pathway</i>	0.030	<i>CDC42;PLCB3;CAMK2D;PLA2G4C;PTK2B;ELK1;CAMK2G</i>
	<i>Phagosome</i>	0.030	<i>ITGB1;DYNC1H1;CLEC4M;LAMP1;ATP6AP1;STX7;HLA-B;HLA-DRB1;HLA-DQA1;THBS3</i>
	<i>Epstein-Barr virus infection</i>	0.034	<i>ATF2;HDAC5;XPO1;PIK3CA;PSMD14;POLR2B;GTF2B;CSNK2A2;HLA-B;HLA-DRB1;HLA-DQA1;POLR2L</i>
	<i>Allograft rejection</i>	0.035	<i>HLA-B;IL2;HLA-DRB1;HLA-DQA1</i>
	<i>Glycerophospholipid metabolism</i>	0.036	<i>GNPAT;PLA2G4C;CHAT;CRLS1;LPIN3;AGPAT3;CDS2</i>
	<i>Pathways in cancer</i>	0.038	<i>ITGB1;TCF7L2;TCF7L1;CSF3R;HSP90AB1;ITGA3;F2R;PDGFA;STK4;PML;FGF5;CDC42;GNA13;PLCB3;RXRA;TFG;PIK3CA;CASP3;TCEB1;FADD</i>
	<i>Leukocyte transendothelial migration</i>	0.040	<i>RAP1B;CDC42;ITGB1;PIK3CA;TXK;PTK2B;CLDN2;BCAR1</i>
	<i>Antigen processing and presentation</i>	0.040	<i>HSP90AB1;HLA-B;KIR3DL2;B2M;HLA-DQA1;HLA-DRB1</i>
	<i>Tryptophan metabolism</i>	0.041	<i>HADHA;ALDH2;AFMID;ALDH9A1</i>
	<i>Pyruvate metabolism</i>	0.041	<i>LDHB;ALDH2;GLO1;ALDH9A1</i>
	<i>Histidine metabolism</i>	0.042	<i>ALDH2;CNBP2;ALDH9A1</i>
	<i>Glycerolipid metabolism</i>	0.043	<i>LIPC;ALDH2;LPIN3;AGPAT3;ALDH9A1</i>
<i>Turquoise</i>	<i>Protein processing in endoplasmic reticulum</i>	0.000	<i>UBE2D4;NGLY1;ERO1L;CAPN1;SEC61B;PDIA3;HSP90AA1;SSR4;YOD1;TRAF2;RAD23B;CKAP4;DNAJC3;DAD1;DDIT3;ERP29;PLAA;PREB;MAN1B1;SAR1A;SAR1B;SEL1L;RPN1;DERL2;UBE2J2;RNF5;HSPBP1;HSP90B1;UBE2J1;DNAJB2;ATXN3;BAG2;BAG1;UBQLN1;SSR1;BCAP31;HSPA8;MBTPS1;SEC24B;XBP1;SEC24A;HSPA5;EDEM2;NSFL1C;STT3A;HYOU1;STT3B;CALR;MBTPS2</i>

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>Metabolic pathways</i>	0.000	<p>AHCYL1;ACSM3;UXS1;XYLT2;EP RS;ENO3;IL4I1;NDST1;MLYCD;AT P6V1E1;SEPHS1;SEPHS2;GLUL;A LG8;ACOT8;ALG6;ACSL1;GLCE;O XSM;ALG3;CEL;HPRT1;ALG10B;C ES1;CDS1;MTMR1;MAOB;SHMT1; RPN1;ATP5I;ATP5H;PLA2G6;LIAS; GNS;PTDSS2;UGCG;SGPL1;CYP2 B6;LTA4H;PLCG1;ARSB;ST3GAL2 ;ATP6V1C2;PCK2;PMM1;CHKA;SL C33A1;MTHFS;PYCR2;CYP4F8;P RDX6;ACADSB;PAFAH2;DHFR;CY P2C9;EXT2;GLUD1;UCK2;AKR1B1 0;QARS;CYP2S1;PEMT;ALDH18A 1;MOCS2;ACY1;GBE1;QPRT;AKR 1B1;ATIC;IMPA1;IMPA2;GUK1;HY AL4;PHGDH;ACADS;ARG2;ARG1; GAA;AKR1A1;NME1;AGPS;RGN;A LDOB;PLCB1;GART;MGLL;PIGH;D TYMK;ECHS1;OAT;AHCY;INPP1;P RIM1;HAAO;ATP5A1;ODC1;PAPS S2;NT5E;MTHFD1L;HSD17B2;GK; GCH1;MDH2;DHRS3;ALDH4A1;DH RS9;NDUFAB1;PLCD1;NDUFA13; CHPF;CPOX;COMT;GALNT10;TM7 SF2;NAGLU;ACAA1;NUDT12;DGA T2;HGD;ALG14;SDHC;ALG11;COX 6B2;SDHA;SDHB;ALPPL2;ATP5J2; ALDH3A2;ACLY;ALDH3A1;DGKQ; DNMT1;NDUFB10;CYP2C19;AK7; PLD3;ACAT1;PMVK;PGK1;MPST; CYP2J2;PDHA1;POLR3GL;EPHX2; GOT2;GFPT1;B3GALT6;DHCR24; COQ7;COQ6;DAK;COQ5;NAT2;ST T3A;STT3B;PNLIPRP2;ACADVL;F TCD;PIK3C2B;SPR;CYP2R1;ME3; A4GALT;GUSB;PGM1;HIBCH;IDH3 A;BCKDHA;DNMT3A;AMT;PCCA;T ST;DAD1;PCCB;NDUFS4;POLR1C ;ALDH1A1;POLR1E;PIK3C3;MVD; MAN1B1;ST6GALNAC4;ASPA;B4G ALT7;RBKS;B4GALT4;NDUFB6;LA MA3;NDUFB1;HSD17B12;HMGCR; AGPAT1;HMGCL;POLD4;CBS;CDI PT;NDUFA8;POLR2J2;IDH3G;B3G AT3;NDUFA2;PLA2G2A;MBOAT1; GSTZ1;POLE4;SUCLA2;PC;GNPD A2;POLR3C;IMPDH1;POLR3D;IMP DH2;CTH;GALM;POLR3H;SCLY;P FKM</p>
---------------------------	-------	--

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>Viral carcinogenesis</i>	0.000	YWHAE;TRADD;HDAC11;CHD4;CREB3L4;CREB3L1;CREB3L2;EP300;RAC1;PRKACB;YWHAH;MRPS18B;TRAF2;ACTN4;HIST2H2BE;CREB3;CREB1;IRF3;CCNE1;VDAC3;TBPL1;HIST1H2BH;HIST1H2BB;HIST1H2BN;GTF2A1;HIST1H2BM;SP100;HIST1H2BI;HDAC1;HPN;PXN;HIST1H2BK;RELA;GTF2E2;HDAC6;C3;NRAS;SCIN;MAPK1;GSN;HIST1H4L;GTF2H3;NFKB1;NFKBIA;DLG1;HIST1H4B;DNAJA3;MDM2;GRB2
<i>SNARE interactions in vesicular transport</i>	0.000	STX17;STX8;GOSR2;GOSR1;BET1L;VTI1A;STX5;STX4;STX3;SEC22B;STX2;VAMP5;BET1;VTI1B
<i>Peroxisome</i>	0.000	PEX16;GSTK1;ACOT8;ACSL1;EPHX2;PEX13;PEX14;PEX7;HMGCL;PRDX5;PEX3;PEX6;PRDX1;PMVK;CAT;AGPS;NUDT19;MLYCD;MPV17;PEX11G;ACAA1;DECR2;PAOX;NUDT12
<i>Alcoholism</i>	0.000	HIST1H2BN;HIST1H2BM;MAOB;HIST1H2BI;HDAC1;HDAC11;HIST1H2BK;HIST2H2AB;CALML3;HDAC6;PPP1CC;NRAS;GNG4;CREB3L4;H2AFJ;CREB3L1;CREB3L2;MAPK1;HIST1H3C;HIST1H2AC;MAP2K1;HIST1H4L;H3F3A;HIST1H2AJ;H2AFX;H2AFV;GNG12;HIST2H2BE;CREB3;HIST1H4B;CREB1;GNB1;PPP1R1B;GNAS;GRB2;HIST1H2BH;CALM1;SLC29A1;HIST1H2BB;CALM2
<i>Prostate cancer</i>	0.000	HSP90AA1;MAP2K1;MAP2K2;PDPK1;PDGFB;RELA;EGFR;NFKB1;HSP90B1;NFKBIA;CREB3;NRAS;CREB1;CREB3L4;CCNE1;CREB3L1;CREB3L2;ERBB2;PDGFC;MDM2;EP300;MAPK1;GRB2;NKX3-1
<i>Vasopressin-regulated water reabsorption</i>	0.000	NSF;RAB5B;RAB5C;DCTN2;ADCY6;CREB3;CREB1;ADCY9;CREB3L4;CREB3L1;CREB3L2;GNAS;STX4;PRKACB;DYNC1I1
<i>Carbon metabolism</i>	0.000	ECHS1;SHMT1;ENO3;ACAT1;PGK1;ME3;PHGDH;ACADS;HIBCH;IDH3A;PDHA1;IDH3G;MDH2;GOT2;AMT;SDHC;SDHA;SDHB;DAK;GLUD1;PC;SUCLA2;PCCA;PCCB;CAT;RGN;ALDOB;PFKM

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>Legionellosis</i>	0.000	HSPA8;SAR1A;SAR1B;RAB1B;BNIP3;CXCL1;CXCL3;CXCL2;RELA;NFKB1;NFKBIA;C3;IL1B;HSF1;SEC22B;MYD88;TLR2
<i>Epstein-Barr virus infection</i>	0.000	YWHAH;PSMD11;HDAC1;TRADD;AKAP8L;RELA;GTF2E2;RELB;PSMD7;IRAK1;MYC;PSMD2;EP300;RIPK1;PLCG1;CD58;PRKACB;MAP2K6;YWHAH;NUP214;MAP2K3;HSPA8;POLR2J2;CSNK2A1;POLR3GL;ENTPD8;TRAF2;NFKB1;NFKBIA;PSMC5;PSMC6;IRF3;PSMC4;POLR3C;POLR3D;POLR3E;CSNK2B;POLR1C;MDM2;TBPL1;POLR3H;VIM;RAN
<i>Adherens junction</i>	0.000	PTPN1;YES1;CSNK2A1;CTNND1;ACTN4;IQGAP1;SORBS1;PTPRF;EGFR;ACTB;TGFB2;ERBB2;CSNK2B;EP300;MAPK1;FYN;RAC1;PVR13;PVRL2;VCL
<i>Citrate cycle (TCA cycle)</i>	0.001	ACLY;PDHA1;PC;SUCLA2;IDH3G;MDH2;SDHC;SDHA;SDHB;PCK2;IDH3A
<i>Shigellosis</i>	0.001	ROCK2;ARPC1B;ARPC4;U2AF1;ARPC5;RELA;ACTB;NFKB1;NFKBIA;CTTN;ARPC2;ARPC3;MAPK1;ELMO3;RAC1;CRK;VCL;ATG5
<i>TNF signaling pathway</i>	0.001	TRADD;CXCL1;CXCL3;CXCL2;RELA;CREB3L4;CASP10;CREB3L1;CREB3L2;MAPK1;RIPK1;DNM1L;MAP2K6;MAP2K3;MAP2K1;JAG1;TRAF2;FOS;NFKB1;NFKBIA;CREB3;ITCH;CREB1;IL1B;FAS;BIRC3
<i>Estrogen signaling pathway</i>	0.001	HSPA8;MAP2K1;HSP90AA1;MAP2K2;CALML3;FOS;EGFR;ADCY6;HSP90B1;CREB3;NRAS;ADCY9;CREB1;CREB3L4;SP1;CREB3L1;CREB3L2;GNAS;MAPK1;GRB2;CALM1;P1CB1;CALM2;PRKACB
<i>RNA degradation</i>	0.001	BTG2;BTG1;WDR61;TOB1;ENO3;LSM3;EXOSC7;EDC4;CNOT6;EXOSC10;EXOSC4;LSM6;CNOT2;XRN2;RQCD1;SKIV2L;EXOSC2;PFKM;DCP2;SKIV2L2
<i>Protein export</i>	0.001	OXA1L;SPCS3;IMMP2L;HSPA5;SRP72;SRPR;SRPRB;SRP68;SEC61B

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>Nucleotide excision repair</i>	0.001	<i>RFC4;PCNA;RFC1;RPA1;RPA2;GTF2H3;RAD23B;GTF2H5;POLE4;POLD4;ERCC3;ERCC1;ERCC8;ERCC5</i>
<i>mRNA surveillance pathway</i>	0.001	<i>UPF2;UPF1;CPSF4;PPP2R5B;PPP2R5D;PPP2R2A;SMG7;GSPT2;SMG5;GSPT1;NXF1;PPP1CC;NUDT21;PABPN1;FIP1L1;PPP2R1B;PPP2R1A;CSTF1;PAPOLA;ETF1;RNPS1;PELO</i>
<i>Fatty acid metabolism</i>	0.002	<i>CPT1A;ACADVL;ECHS1;ELOVL5;ACSL1;OXSM;HSD17B12;CPT1B;ACADSB;ACAT1;CPT2;SCD;ACAA1;ACADS</i>
<i>Salmonella infection</i>	0.002	<i>ROCK2;ARPC1B;IFNGR2;CXCL1;ARPC4;FOS;ARPC5;CXCL3;CXCL2;RELA;ACTB;NFKB1;KLC4;ARPC2;ARPC3;IL1B;CCL4;MAPK1;RAC1;DYNC111;MYD88</i>
<i>Glycosaminoglycan biosynthesis - heparan sulfate / heparin</i>	0.002	<i>EXT2;NDST1;B3GAT3;GLCE;XYLT2;B3GALT6;HS6ST1;B4GALT7;HS3ST1</i>
<i>One carbon pool by folate</i>	0.002	<i>DHFR;ATIC;MTHFD1L;SHMT1;AMT;MTHFS;GART;FTCD</i>
<i>AMPK signaling pathway</i>	0.002	<i>PFKFB2;IRS1;PPP2R2A;HMGCR;ADIPOR2;GYS1;PPP2R1B;PPP2R1A;CREB3L4;CREB3L1;CREB3L2;EIF4EBP1;MLYCD;PCK2;SREBF1;CPT1A;PDPK1;PPP2R5B;TSC2;PPP2R5D;CPT1B;CREB3;TBC1D1;CREB1;SCD;ULK1;PFKM</i>
<i>Amphetamine addiction</i>	0.003	<i>MAOB;HDAC1;CALML3;FOS;CACNA1C;PPP1CC;CREB3;PPP3R1;CREB1;CREB3L4;CREB3L1;CREB3L2;PPP1R1B;GNAS;CALM1;CALM2;PRKACB</i>
<i>Selenocompound metabolism</i>	0.003	<i>TXNRD2;TXNRD1;CTH;SCLY;SEPHS1;SEPHS2;PAPSS2</i>
<i>Insulin signaling pathway</i>	0.003	<i>IRS1;CALML3;CBL;PTPRF;SOCS2;PPP1CC;GYS1;NRAS;MKNK2;EIF4EBP1;FLOT2;MAPK1;PRKACB;SH2B2;PCK2;PTPN1;SREBF1;MAP2K1;MAP2K2;PDPK1;TSC2;PHKB;SORBS1;GRB2;CALM1;EIF4E2;CALM2;CRK;RHOQ</i>

continua



**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>DNA replication</i>	0.003	<i>POLE4;POLD4;RFC4;PCNA;RFC1;MCM7;PRIM1;RPA1;RPA2;MCM4;MCM6</i>
<i>Insulin resistance</i>	0.003	<i>MLXIP;PTPN1;MGEA5;SREBF1;SLC27A1;CPT1A;IRS1;PDPK1;GFPT1;SLC2A1;CPT1B;RELA;NFKB1;PTPRF;NFKBIA;PPP1CC;CREB3;GYS1;CREB1;CREB3L4;CREB3L1;CREB3L2;OGT;PCK2</i>
<i>Biosynthesis of amino acids</i>	0.003	<i>ARG2;ACY1;ARG1;IDH3G;SHMT1;GOT2;PYCR2;ENO3;PC;CBS;CTH;PGK1;PHGDH;ALDH18A1;ALDOB;PFKM;GLUL;IDH3A</i>
<i>Dopaminergic synapse</i>	0.004	<i>MAOB;CALML3;PPP2R2A;CACNA1C;COMT;ARNTL;PPP1CC;PPP2R1B;PPP2R1A;GNG4;CREB3L4;CREB3L1;CREB3L2;PRKACB;DRD5;PPP2R5B;PPP2R5D;FOS;GNG12;CREB3;CREB1;GNB1;PPP1R1B;GNAS;CALM1;PLCB1;CALM2</i>
<i>Sulfur relay system</i>	0.004	<i>MPST;TRMU;MOCS2;MOCS3;TST</i>
<i>Valine, leucine and isoleucine degradation</i>	0.004	<i>BCKDHA;ECHS1;ACADSB;AACS;ACAT1;ALDH3A2;IL4I1;HMGCL;PCA;PCCB;ACAA1;ACADS;HIBCH</i>
<i>Thyroid hormone signaling pathway</i>	0.005	<i>PFKFB2;NOTCH3;THRA;HDAC1;SLC2A1;ATP2A2;ATP1A1;ACTB;RXRB;NRAS;MYC;EP300;MAPK1;PLCG1;PRKACB;NCOA1;MAP2K1;MAP2K2;STAT1;PDPK1;TSC2;FXYD2;MDM2;PLCB1;PLCD1</i>
<i>Ubiquinone and other terpenoid-quinone biosynthesis</i>	0.007	<i>NQO1;VKORC1L1;COQ7;COQ6;COQ5</i>
<i>Endocytosis</i>	0.007	<i>RAB5B;RAB5C;TSG101;TFRC;ARPC1B;VPS4B;KIAA1033;SNX12;CBLL;EGFR;PARD6B;ADRBK1;CHMP1B;CLTCL1;FAM21C;CCDC53;SNX5;AP2M1;SNX6;ARFGEF1;HSPA8;IQSEC1;CAV1;ZFYVE20;KIAA0196;ARPC4;SNF8;ARPC5;ARFGAP3;DNM1;TGFB2;RUFY1;ITCH;DAB2;ARPC2;BIN1;ARPC3;RAB35;CHMP2A;MDM2;CHMP4B;STAMPB;FOLR3;SMAP1;VPS28;ARF6</i>

continua

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>Alzheimer's disease</i>	0.008	APP;NDUFA13;LRP1;NDUFB10;NDUFB6;ATP5A1;PSEN2;NDUFB1;ATP2A2;CALML3;CACNA1C;IDE;ATP5H;APH1A;PPP3R1;MAPK1;CAPN1;NDUFA8;NDUFA2;SDHC;SDHA;COX6B2;SDHB;BACE1;CDK5;IL1B;NDUFS4;NDUFAB1;FAS;CALM1;PLCB1;CALM2
<i>Melanogenesis</i>	0.010	MAP2K1;MAP2K2;FZD5;CALML3;ADCY6;CREB3;NRAS;KITLG;ADCY9;CREB1;CREB3L4;CREB3L1;CREB3L2;GNAS;DVL2;EP300;MAPK1;CALM1;PLCB1;CALM2;PRKACB
<i>Glucagon signaling pathway</i>	0.011	CPT1A;PDHA1;SMEK2;SLC2A1;CALML3;PHKB;CPT1B;CREB3;GYS1;PPP3R1;CREB1;CREB3L4;CREB3L1;CREB3L2;GNAS;EP300;CALM1;PLCB1;CALM2;PRKACB;PCK2
<i>Thyroid hormone synthesis</i>	0.012	HSPA5;GPX3;GSR;ATP1A1;ADCY6;HSP90B1;CREB3;ADCY9;CREB1;CREB3L4;FXVD2;CREB3L1;CREB3L2;GNAS;PLCB1;PRKACB
<i>Apoptosis</i>	0.012	TUBAL3;TRADD;HTRA2;RELA;ACTB;LMNB2;NRAS;CASP10;LMNA;CASP2;MAPK1;RIPK1;CAPN1;CTSD;CTSC;MCL1;MAP2K1;ENDOG;MAP2K2;PDPK1;TRAF2;FOS;NFKB1;NFKBIA;DDIT3;FAS;BIRC3
<i>cGMP-PKG signaling pathway</i>	0.013	ROCK2;IRS1;ATP2A2;CALML3;ATP1A1;CACNA1C;ADCY6;PPP1CC;PPP3R1;RGS2;CREB3L4;CREB3L1;CREB3L2;GNA12;MAPK1;MEF2D;MEF2A;MAP2K1;MAP2K2;MEF2B;ATP2B4;CREB3;ADCY9;CREB1;FXVD2;VDAC3;GTF2IRD1;VDAC1;CALM1;PLCB1;CALM2
<i>Renal cell carcinoma</i>	0.013	MAP2K1;EGLN2;MAP2K2;EPAS1;CUL2;SLC2A1;PDGFB;NRAS;EP300;MAPK1;GRB2;VHL;RAC1;PAK2;CRK
<i>Bacterial invasion of epithelial cells</i>	0.013	CAV1;ARPC1B;PXN;ILK;ARPC4;ARPC5;CBL;DNM1;ACTB;CTTN;ARPC2;ARPC3;CLTCL1;ELMO3;RAC1;CRK;VCL
<i>HIF-1 signaling pathway</i>	0.014	MAP2K1;MAP2K2;PDHA1;EGLN2;TFRC;IFNGR2;CUL2;SLC2A1;ENO3;RELA;EGFR;NFKB1;MKNK2;ERBB2;EIF4EBP1;PGK1;EP300;MAPK1;PLCG1;VHL;EIF4E2

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>Cocaine addiction</i>	0.014	CREB3;CREB1;CDK5;MAOB;CREB3L4;CREB3L1;CREB3L2;PPP1R1B;GNAS;PRKACB;RELA;NFKB1
<i>GnRH signaling pathway</i>	0.015	MAP3K2;MAP2K3;MAP2K1;MAP2K2;CALML3;CACNA1C;EGFR;ADCY6;NRAS;ADCY9;GNAS;MAPK1;GRB2;CALM1;PLCB1;CALM2;PRKACB;MAP3K4;MAP2K6
<i>Pathogenic Escherichia coli infection</i>	0.015	TUBAL3;ROCK2;TUBB;ARPC1B;ARPC4;ARPC5;ACTB;OCLN;CTTN;ARPC2;ARPC3;NCK2;FYN
<i>Lysosome</i>	0.015	ARSA;CD63;GAA;SLC11A2;AP1B1;MCOLN1;GNS;GGA2;GGA3;NAGLU;GM2A;NPC2;AP1G1;LAMP3;CLTCL1;NEU1;SLC17A5;TPP1;MAN2B1;AP1S3;GUSB;CTSD;ARSB;CTSC
<i>Fatty acid degradation</i>	0.016	ALDH3A2;CPT1A;ACADVL;ECHS1;CPT2;ACSL1;ACAA1;CPT1B;ACADSB;ACADS;ACAT1
<i>Proteasome</i>	0.016	PSMC5;PSMB4;PSMC6;PSMD11;PSMD7;PSMC4;PSME3;PSMD2;PSME4;PSMA7;PSMB10
<i>Glyoxylate and dicarboxylate metabolism</i>	0.017	PCCA;MDH2;SHMT1;PCCB;CAT;AMT;GLUL;ACAT1
<i>Mismatch repair</i>	0.018	POLD4;PCNA;RFC4;RFC1;MSH3;RPA1;RPA2
<i>Fc gamma R-mediated phagocytosis</i>	0.018	MAP2K1;GSN;LIMK2;ARPC1B;ARPC4;ARPC5;PLA2G6;VAV2;SCIN;ARPC2;ARPC3;BIN1;CFL1;MAPK1;PLCG1;RAC1;CRK;LAT;ARF6
<i>Cysteine and methionine metabolism</i>	0.019	MPST;IL4I1;DNMT1;AHCYL1;AHCY;TST;MDH2;CBS;CTH;GOT2;DNMT3A
<i>Pertussis</i>	0.019	C4BPA;CALML3;C4BPB;FOS;RELA;NFKB1;C3;IRF3;IRAK1;IL1B;CFL1;IRF8;MAPK1;CALM1;CALM2;MYD88
<i>PPAR signaling pathway</i>	0.020	SLC27A1;CPT1A;GK;ACSL1;PDPK1;ILK;DBI;SORBS1;CPT1B;RXRB;CPT2;SCD;FABP6;ACAA1;PCK2
<i>Longevity regulating pathway - mammal</i>	0.020	IRS1;TSC2;ADIPOR2;RELA;ADCY6;NFKB1;CREB3;NRAS;ADCY9;CREB1;CREB3L4;CREB3L1;CREB3L2;CAT;EIF4EBP1;ULK1;EIF4E2;PRKACB;ATG5

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>Hepatitis B</i>	0.020	PCNA;RELA;NRAS;CREB3L4;CAS P10;CREB3L1;MYC;CREB3L2;EP300;MAPK1;MAP2K1;MAP2K2;STAT1;FOS;HSPG2;NFKB1;NFKBIA;CREB3;CREB1;IRF3;CCNE1;VDAC3;FAS;GRB2;MYD88;TLR3;TLR2
<i>Thyroid cancer</i>	0.021	RXRB;MAP2K1;NRAS;MAP2K2;MYC;TPR;CCDC6;MAPK1
<i>Choline metabolism in cancer</i>	0.022	MAP2K1;MAP2K2;SLC22A5;SLC44A4;CHKA;SLC44A1;PDPK1;PDGF B;TSC2;FOS;EGFR;NRAS;SP1;DGKQ;PDGFC;EIF4EBP1;MAPK1;GRB2;PLCG1;RAC1
<i>Folate biosynthesis</i>	0.022	DHFR;MOCS2;GCH1;SPR;ALPPL2
<i>Endometrial cancer</i>	0.023	APC2;NRAS;MAP2K1;MAP2K2;PDPK1;MYC;ERBB2;AXIN1;ILK;MAPK1;GRB2;EGFR
<i>Cytosolic DNA-sensing pathway</i>	0.023	POLR3GL;ADAR;RELA;NFKB1;NFKBIA;IRF3;POLR3C;IL1B;POLR3D;POLR3E;POLR1C;CCL4;POLR3H;RIPK1
<i>Adrenergic signaling in cardiomyocytes</i>	0.024	CREM;ATP2A2;CALML3;PPP2R2A;ATP1A1;CACNA1C;ADCY6;PPP1CC;PPP2R1B;PPP2R1A;CREB3L4;CREB3L1;CREB3L2;MAPK1;PRKACB;TPM4;PPP2R5B;ATP2B4;PPP2R5D;CREB3;ADCY9;CREB1;FXYD2;GNAS;CALM1;PLCB1;CALM2
<i>Glycerolipid metabolism</i>	0.026	PNLIPRP2;DGAT2;GK;AKR1A1;MBOAT1;AKR1B1;CEL;AGPAT1;DAK;ALDH3A2;AKR1B10;DGKQ;MGLL
<i>Purine metabolism</i>	0.026	PRIM1;NUDT2;AK7;FHIT;HDDC3;ADCY6;PAPSS2;POLD4;NT5E;ATIC;GUK1;ENPP4;PGM1;ENTPD2;POLR2J2;POLR3GL;ENTPD6;ENTPD8;NME1;POLE4;ADCY9;POLR3C;IMPDH1;POLR3D;IMPDH2;POLR3E;POLR1C;POLR1E;POLR3H;HPRT1;GART
<i>Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate</i>	0.030	CHPF;B3GAT3;XYLT2;B3GALT6;B4GALT7;CHST3
<i>Arginine biosynthesis</i>	0.030	GLUD1;ARG2;ACY1;ARG1;GOT2;GLUL
<i>Pyrimidine metabolism</i>	0.032	DTYMK;POLR2J2;POLR3GL;TXNRD2;TXNRD1;PRIM1;ENTPD6;ENTPD8;NUDT2;NME1;POLE4;POLD4;NT5E;UCK2;POLR3C;POLR3D;POLR3E;POLR1C;POLR1E;POLR3H

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>N-Glycan biosynthesis</i>	0.034	ALG8;ALG6;DAD1;RPN1;ALG14;STT3A;ALG3;ALG11;STT3B;MAN1B1;ALG10B
<i>Toll-like receptor signaling pathway</i>	0.035	MAP2K3;MAP2K1;MAP2K2;STAT1;FOS;RELA;NFKB1;NFKBIA;IRF3;IRAK1;TOLLIP;IL1B;CCL4;MAPK1;RIPK1;RAC1;MYD88;TLR3;MAP2K6;TLR2
<i>NF-kappa B signaling pathway</i>	0.035	CSNK2A1;TRADD;TRAF2;CXCL2;RELA;NFKB1;RELB;NFKBIA;IRAK1;PLAU;IL1B;CSNK2B;CCL4;RIPK1;PLCG1;LAT;MYD88;BIRC3
<i>Propanoate metabolism</i>	0.037	BCKDHA;SUCLA2;ECHS1;PCCA;PCCB;MLYCD;HIBCH;ACAT1
<i>RNA polymerase</i>	0.037	POLR2J2;POLR3GL;POLR3C;POLR3D;POLR3E;POLR1C;POLR1E;POLR3H
<i>Axon guidance</i>	0.038	EPHB6;SEMA6A;ROCK2;LIMK2;SEMA4C;NTN4;EFNA4;ROBO1;EFNB2;NRAS;PPP3R1;CDK5;RASA1;CFL1;NCK2;MAPK1;FYN;RAC1;EPHB2;PAK2;NGEF;EPHB4;PLXNA3
<i>Arginine and proline metabolism</i>	0.039	ALDH4A1;ALDH3A2;ARG2;OAT;MAOB;SMOX;ARG1;ODC1;GOT2;PYCR2;ALDH18A1
<i>Dorso-ventral axis formation</i>	0.042	PIWIL2;NOTCH3;MAP2K1;PIWIL4;MAPK1;GRB2;EGFR
<i>Pathways in cancer</i>	0.042	SLC2A1;LAMC2;CKS1B;MYC;EP300;RAC1;PRKACB;NKX3-1;APC2;MAP2K1;HSP90AA1;RALBP1;MAP2K2;AXIN1;TRAF2;FOS;TGFB2;ADCY9;TRAF4;MSH3;CCNE1;CKS2;PLCB1;CRK;BIRC3;ROCK2;HDAC1;EPAS1;CUL2;LAMA3;PDGFB;CBL;CSF2RA;ADCY6;EGFR;RELA;HSP90B1;RXRB;NRAS;GNG4;TPR;ERBB2;GNA12;DVL2;MAPK1;PLCG1;VHL;EGLN2;FZD5;STAT1;GNG12;NFKB1;NFKBIA;KITLG;BMP2;GNB1;CCDC6;MDM2;GNAS;FAS;GRB2
<i>Vibrio cholerae infection</i>	0.044	ADCY9;KDELR1;GNAS;SEC61B;PLCG1;ATP6V1E1;PRKACB;ACTB;TJP2;ERO1L;ATP6V1C2
<i>Basal transcription factors</i>	0.045	GTF2A1;ERCC3;TAF11;TAF9;TBP1;GTF2IRD1;TAF6;GTF2H3;GTF2H5;GTF2E2
<i>Brown</i>	0.000	GSK3B;EXOC7;ARAF;CBL;IRS2;PPP1CA;PPP1R3B;RPS6KB2;FLOT1;TRIP10;AKT1;CALM3;KRAS

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>Chronic myeloid leukemia</i>	0.000	CDKN1A;CDK6;ARAF;CBLC;AKT1;KRAS;IKBK;RUNX1;BCL2L1
<i>Acute myeloid leukemia</i>	0.001	JUP;RPS6KB2;ARAF;AKT1;KRAS;IKBK;RUNX1
<i>Hepatitis C</i>	0.001	GSK3B;CDKN1A;OAS3;ARAF;IRF7;AKT1;KRAS;TYK2;IKBK;CLDN1;PIAS1
<i>Pancreatic cancer</i>	0.001	RALA;CDK6;ARAF;AKT1;KRAS;IKBK;BCL2L1
<i>Proteoglycans in cancer</i>	0.002	CDKN1A;FZD7;ARAF;CBLC;HIF1A;PPP1CA;VTN;RRAS;RPS6KB2;AKT1;PTPN6;KRAS;MET
<i>Melanoma</i>	0.002	CDKN1A;CDK6;CDH1;ARAF;AKT1;KRAS;MET
<i>Lysosome</i>	0.004	AP4M1;GNPTG;ATP6V0B;LAMP2;GBA;CTSH;LIPA;IGF2R;AP1M2
<i>Proteasome</i>	0.005	PSMB6;PSMA3;PSMC3;PSMD3;PSMF1
<i>Non-alcoholic fatty liver disease (NAFLD)</i>	0.005	GSK3B;CASP7;NDUFS8;NDUFA5;NDUFB11;AKT1;UQCRFS1;IRS2;ADIPOR1;UQCRH
<i>Longevity regulating pathway - multiple species</i>	0.006	CLPB;RPS6KB2;AKT1;KRAS;IRS2;SOD1
<i>Glioma</i>	0.006	CDKN1A;CDK6;ARAF;AKT1;CALM3;KRAS
<i>Metabolic pathways</i>	0.006	LAMA5;IDI1;ISYNA1;GPAA1;NDUFB11;MAOA;GBA;PDXP;ENO1;MTMR4;UQCRH;SMPD3;NT5C;CHSY1;UROD;FLAD1;MAN2A1;IDH3B;HYI;PGM2;UQCRFS1;MGAT1;ACADM;DLAT;GGT1;COX10;PSPH;ATP6V1F;GAMT;ATP6V0B;PCYT2;MDH1;NDUFA5;PTGES2;DCXR;NME4;ALG12;PISD;SQLE;DPM2;GALE;NDUFS8;CPS1;NME6;SMS;PNPLA2
<i>HTLV-I infection</i>	0.006	GSK3B;VAC14;CDKN1A;TBP;FZD7;IL2RG;FOSL1;CCND2;RRAS;AKT1;KRAS;IKBK;ANAPC1;BCL2L1
<i>ErbB signaling pathway</i>	0.007	GSK3B;CDKN1A;RPS6KB2;ARAF;CBLC;AKT1;KRAS
<i>Oxidative phosphorylation</i>	0.007	ATP6V0B;NDUFS8;PPA2;NDUFA5;NDUFB11;UQCRFS1;UQCRH;COX10;ATP6V1F
<i>Apoptosis</i>	0.009	DAXX;CASP7;DFFA;AKT1;CTSH;KRAS;BAK1;IKBK;BCL2L1
<i>Endometrial cancer</i>	0.010	GSK3B;CDH1;ARAF;AKT1;KRAS
<i>B cell receptor signaling pathway</i>	0.010	LYN;GSK3B;AKT1;PTPN6;KRAS;IKBK

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>Pathways in cancer</i>	0.013	GSK3B;LAMA5;RALA;CDKN1A;JUP;TPM3;FZD7;ARAF;CBL;HIF1A;RUNX1;CDK6;CDH1;AKT1;KRAS;IKBK;MET;BCL2L1
<i>Cell cycle</i>	0.014	CDKN1C;GSK3B;WEE1;CDKN1A;CCND2;CDK6;SFN;ANAPC1
<i>PI3K-Akt signaling pathway</i>	0.014	GSK3B;LAMA5;CDKN1A;PPP2R3A;IL2RG;VTN;CCND2;CDK6;COL6A2;CDC37;RPS6KB2;AKT1;KRAS;IKBK;MET;BCL2L1
<i>Glycine, serine and threonine metabolism</i>	0.018	GAMT;MAOA;GCAT;PSPH
<i>Jak-STAT signaling pathway</i>	0.020	PIAS4;CDKN1A;CCND2;AKT1;PTPN6;TYK2;IL2RG;PIAS1;BCL2L1
<i>Bladder cancer</i>	0.020	CDKN1A;CDH1;ARAF;KRAS
<i>Herpes simplex infection</i>	0.020	TAF6L;DAXX;TBP;CDC34;OAS3;IRF7;TYK2;IKBK;PPP1CA;SRPK1
<i>FoxO signaling pathway</i>	0.020	CDKN1A;CCND2;ARAF;AKT1;HOMER3;IRS2;KRAS;NLK
<i>Insulin resistance</i>	0.021	GSK3B;PPP1R3B;NR1H2;RPS6KB2;AKT1;IRS2;PPP1CA
<i>Measles</i>	0.023	GSK3B;CCND2;CDK6;OAS3;IRF7;AKT1;TYK2;IL2RG
<i>Carbon metabolism</i>	0.025	CPS1;MDH1;IDH3B;ENO1;ACADM;DLAT;PSPH
<i>Prostate cancer</i>	0.026	GSK3B;CDKN1A;ARAF;AKT1;KRAS;IKBK
<i>Renal cell carcinoma</i>	0.026	ARAF;AKT1;KRAS;MET;HIF1A
<i>Central carbon metabolism in cancer</i>	0.027	AKT1;KRAS;SLC1A5;MET;HIF1A
<i>Alzheimer's disease</i>	0.028	GSK3B;CASP7;NDUFS8;NDUFA5;NDUFB11;UQCRRS1;CALM3;UQCRRH;CDK5R1
<i>Epithelial cell signaling in Helicobacter pylori infection</i>	0.029	LYN;ATP6V0B;IKBK;MET;ATP6V1F
<i>Longevity regulating pathway - mammal</i>	0.032	EHMT2;RPS6KB2;AKT1;KRAS;IRS2;ADIPOR1
<i>N-Glycan biosynthesis</i>	0.035	DPM2;MAN2A1;ALG12;MGAT1
<i>Thyroid cancer</i>	0.036	TPM3;CDH1;KRAS
<i>Viral carcinogenesis</i>	0.037	LYN;VAC14;CDKN1A;CCND2;CDK6;TBP;IRF7;KRAS;IKBK;BAK1
<i>Citrate cycle (TCA cycle)</i>	0.040	MDH1;IDH3B;DLAT
<i>Amyotrophic lateral sclerosis (ALS)</i>	0.040	DAXX;TOMM40;SOD1;BCL2L1
<i>Glutathione metabolism</i>	0.043	GPX2;SMS;MGST2;GGT1

*continua*

**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472** *continuação*

<i>Pink</i>	<i>Ubiquitin mediated proteolysis</i>	0.002	<i>PIAS3;UBE2R2;UBE2D2;UBE2NL;TCEB2;UBE4A;DDB2</i>
	<i>Metabolic pathways</i>	0.002	<i>ST6GALNAC1;DDC;NDUFB5;RPE;FUT1;AGPAT2;PAPSS1;EBP;LDHA;POLR2A;POLD1;PNPO;POLR2F;EXTL3;AASS;MCCC2;UGT2B11;IDH1;NDUFC2;ALG1;APRT;MTHFD1;ADI1;ETNK1;PPT2;POLR3K;SLC27A5;ITPA</i>
	<i>Purine metabolism</i>	0.007	<i>POLR2A;POLD1;POLR2F;POLR3K;ITPA;PAPSS1;APRT</i>
	<i>RNA polymerase</i>	0.007	<i>POLR2A;POLR2F;POLR3K</i>
	<i>Colorectal cancer</i>	0.008	<i>CASP9;TGFB1;TCF7;RAC3</i>
	<i>Base excision repair</i>	0.008	<i>POLD1;XRCC1;APEX2</i>
	<i>Pancreatic cancer</i>	0.009	<i>RB1;CASP9;TGFB1;RAC3</i>
	<i>Pyrimidine metabolism</i>	0.010	<i>POLR2A;POLD1;POLR2F;POLR3K;ITPA</i>
	<i>p53 signaling pathway</i>	0.011	<i>CASP9;STEAP3;PTEN;DDB2</i>
	<i>Adherens junction</i>	0.014	<i>PARD3;TCF7;RAC3;MLLT4</i>
	<i>Arrhythmogenic right ventricular cardiomyopathy (ARVC)</i>	0.014	<i>CACNB3;TCF7;CACNA1D;CACNG3</i>
	<i>ABC transporters</i>	0.018	<i>ABCA2;ABCB1;ABCB6</i>
	<i>Phenylalanine metabolism</i>	0.019	<i>DDC;MIF</i>
	<i>Hypertrophic cardiomyopathy (HCM)</i>	0.020	<i>CACNB3;TGFB1;CACNA1D;CACNG3</i>
	<i>Nucleotide excision repair</i>	0.021	<i>POLD1;GTF2H1;DDB2</i>
	<i>RNA transport</i>	0.021	<i>EIF1AX;DDX20;EIF4EBP2;EIF2B1;RANGAP1;EIF4B</i>
	<i>Malaria</i>	0.023	<i>TGFB1;SDC1;HBA2</i>
	<i>Prostate cancer</i>	0.025	<i>RB1;CASP9;TCF7;PTEN</i>
	<i>Dilated cardiomyopathy</i>	0.026	<i>CACNB3;TGFB1;CACNA1D;CACNG3</i>
	<i>Endometrial cancer</i>	0.027	<i>CASP9;TCF7;PTEN</i>
	<i>Pathways in cancer</i>	0.028	<i>RB1;CASP9;FZD2;TGFB1;PLEKHG5;TCF7;PTEN;TCEB2;RAC3;GNAI2</i>
	<i>Renin-angiotensin system</i>	0.033	<i>NLN;AGT</i>
	<i>Histidine metabolism</i>	0.036	<i>DDC;HNMT</i>
	<i>Hepatitis B</i>	0.037	<i>RB1;CASP9;TGFB1;PTEN;DDB2</i>
	<i>Adrenergic signaling in cardiomyocytes</i>	0.038	<i>CACNB3;PPP2R5E;CACNA1D;CACNG3;GNAI2</i>

continua



**Anexo 4 - Enriquecimento funcional com base no banco KEGG dos módulos associados às janelas da rede do grupo Caco-2/Ec472 conclusão**

---

<i>Proteoglycans in cancer</i>	0.043	<i>FZD2;TGFB1;GPC1;SDC1;FLNA;EIF4B</i>
<i>Hippo signaling pathway</i>	0.043	<i>TGFB1;FZD2;PARD3;TCF7;TEAD3</i>

---

## 8 REFERÊNCIAS

Androulakis IP, Yang E, Almon RR. Analysis of time-series gene expression data: methods, challenges, and opportunities. *Annu Rev Biomed Eng.* 2007;9:205-28.

Aoki K, Ogata Y, Shibata D. Approaches for extracting practical information from gene co-expression networks in plant biology. *Plant Cell Physiol.* 2007;48(3):381-90.

Askar M, Faber MS, Frank C, Bernard H, Gilsdorf A, Fruth A, et al. Update on the ongoing outbreak of haemolytic uraemic syndrome due to Shiga toxin-producing *Escherichia coli* (STEC) serotype O104, Germany, May 2011. *Euro Surveill.* 2011;16(22).

Azevedo H, Fujita A, Bando SY, Iamashita P, Moreira-Filho CA. Transcriptional network analysis reveals that AT1 and AT2 angiotensin II receptors are both involved in the regulation of genes essential for glioma progression. *PLoS One.* 2014;9(11):e110934.

Bakovic M, Fullerton MD, Michel V. Metabolic and molecular aspects of ethanolamine phospholipid biosynthesis: the role of CTP:phosphoethanolamine cytidyltransferase (Pcyt2). *Biochem Cell Biol.* 2007;85(3):283-300.

Bando SY, Alegro MC, Amaro E, Silva AV, Castro LH, Wen HT, et al. Hippocampal CA3 transcriptome signature correlates with initial precipitating injury in refractory mesial temporal lobe epilepsy. *PLoS One.* 2011;6(10):e26268.

Bando SY, Iamashita P, Guth BC, dos Santos LF, Fujita A, Abe CM, Ambrosio MRS, Ferreira LR, Moreira-Filho CA. A transcriptional module in HUS-associated STEC induces gene network dysregulation in Caco-2 cells. *Plos One.* In press 2017.

Bando SY, Silva FN, Costa LaF, Silva AV, Pimentel-Silva LR, Castro LH, et al. Complex network analysis of CA3 transcriptome reveals pathogenic and compensatory pathways in refractory temporal lobe epilepsy. *PLoS One.* 2013;8(11):e79913.

Bar-Joseph Z, Gitter A, Simon I. Studying and modelling dynamic biological processes using time-series gene expression data. *Nat Rev Genet.* 2012;13(8):552-64.

- Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet.* 2011;12(1):56-68.
- Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet.* 2004;5(2):101-13.
- Bhattacharya S, Srinivasan K, Abdisalaam S, Su F, Raj P, Dozmorov I, et al. RAD51 interconnects between DNA replication, DNA repair and immunity. *Nucleic Acids Res.* 2017;45(8):4590-605.
- Béjar S, Bouché F, Bouché JP. Cell division inhibition gene *dicB* is regulated by a locus similar to lambdaoid bacteriophage immunity loci. *Mol Gen Genet.* 1988;212(1):11-9.
- Bejar S, Cam K, Bouche JP. Control of cell-division in *Escherichia coli* - dna-sequence of *dica* and of a 2nd gene complementing mutation *dical*, *dicc*. *Nucleic Acids Research.* 1986;14(17):6821-33.
- Bennett KM, Walker SL, Lo DD. Epithelial microvilli establish an electrostatic barrier to microbial adhesion. *Infect Immun.* 2014;82(7):2860-71.
- Benson M, Breitling R. Network theory to understand microarray studies of complex diseases. *Curr Mol Med.* 2006;6(6):695-701.
- Berin MC, Darfeuille-Michaud A, Egan LJ, Miyamoto Y, Kagnoff MF. Role of EHEC O157:H7 virulence factors in the activation of intestinal epithelial cell NF-kappaB and MAP kinase pathways and the upregulated expression of interleukin 8. *Cell Microbiol.* 2002;4(10):635-48.
- Bielaszewska M, Friedrich AW, Aldick T, Schürk-Bulgrin R, Karch H. Shiga toxin activatable by intestinal mucus in *Escherichia coli* isolated from humans: predictor for a severe clinical outcome. *Clin Infect Dis.* 2006;43(9):1160-7.
- Bielaszewska M, Rüter C, Bauwens A, Greune L, Jarosch KA, Steil D, et al. Host cell interactions of outer membrane vesicle-associated virulence factors of enterohemorrhagic *Escherichia coli* O157: Intracellular delivery, trafficking and mechanisms of cell injury. *PLoS Pathog.* 2017;13(2):e1006159.
- Bioconductor open source software for bioinformatics [Internet]. 2003. Available from: <http://www.bioconductor.org/>.
- Blanco M, Blanco JE, Mora A, Dahbi G, Alonso MP, González EA, et al. Serotypes, virulence genes, and intimin types of Shiga toxin (verotoxin)-producing *Escherichia coli* isolates from cattle in Spain and identification of a new intimin variant gene (*eae-xi*). *J Clin Microbiol.* 2004;42(2):645-51.

- Bouwmeester T, Bauch A, Ruffner H, Angrand PO, Bergamini G, Croughton K, et al. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat Cell Biol.* 2004;6(2):97-105.
- Brusa V, Aliverti V, Aliverti F, Ortega EE, de la Torre JH, Linares LH, et al. Shiga toxin-producing *Escherichia coli* in beef retail markets from Argentina. *Frontiers in Cellular and Infection Microbiology.* 2013;3.
- Bugarel M, Beutin L, Martin A, Gill A, Fach P. Micro-array for the identification of Shiga toxin-producing *Escherichia coli* (STEC) seropathotypes associated with Hemorrhagic Colitis and Hemolytic Uremic Syndrome in humans. *Int J Food Microbiol.* 2010;142(3):318-29.
- Buvens G, Piérard D. Virulence profiling and disease association of verocytotoxin-producing *Escherichia coli* O157 and non-O157 isolates in Belgium. *Foodborne Pathog Dis.* 2012;9(6):530-5.
- Carter H, Hofree M, Ideker T. Genotype to phenotype via network analysis. *Current Opinion in Genetics & Development.* 2013;23(6):611-21.
- Celis-Gutierrez J, Boyron M, Walzer T, Pandolfi PP, Jonjić S, Olive D, et al. Dok1 and Dok2 proteins regulate natural killer cell development and function. *EMBO J.* 2014;33(17):1928-40.
- Centro de Vigilância Epidemiológica. *Síndrome hemolítico-urêmica - Normas e Instruções.* São Paulo: CVE; 2002.
- Cerqueira AMF, Tibana A, Guth BEC. High occurrence of shiga-like toxin-producing strains among diarrheagenic *Escherichia coli* isolated from raw beef products in Rio de Janeiro City, Brazil. *Journal of Food Protection.* 1997;60(2):177-80.
- Chambers JE, Dalton LE, Clarke HJ, Malzer E, Dominicus CS, Patel V, et al. Actin dynamics tune the integrated stress response by regulating eukaryotic initiation factor 2 $\alpha$  dephosphorylation. *Elife.* 2015;4.
- Chen L, Yu JH, Lu ZH, Zhang W. E2F2 induction is related to cell proliferation and poor prognosis in non-small cell lung carcinoma. *Int J Clin Exp Pathol.* 2015;8(9):10545-54.
- Choi YJ, Im E, Chung HK, Pothoulakis C, Rhee SH. TRIF mediates Toll-like receptor 5-induced signaling in intestinal epithelial cells. *J Biol Chem.* 2010;285(48):37570-8.
- Chu L-H, Chen B-S. Construction of a cancer-perturbed protein-protein interaction network for discovery of apoptosis drug targets. *Bmc Systems Biology.* 2008;2.

Chuang H-Y, Hofree M, Ideker T. A Decade of Systems Biology. *Annual Review of Cell and Developmental Biology*, Vol 26. 2010;26:721-44.

Clauset A, Shalizi CR, Newman MEJ. Power-law distributions in empirical data. *SIAM Rev.* 2009;51(4),661-703.

Concentric measurements software. Silva FN, Costa L da N. 2013 [Internet]. Available from: <http://cyvision.ifsc.usp.br/concentric/software>.

Crespi A, Bertoni A, Ferrari I, Padovano V, Della Mina P, Berti E, et al. POF1B localizes to desmosomes and regulates cell adhesion in human intestinal and keratinocyte cell lines. *J Invest Dermatol.* 2015;135(1):192-201.

Croxen MA, Finlay BB. Molecular mechanisms of Escherichia coli pathogenicity. *Nature Reviews Microbiology.* 2010;8(1):26-38.

Cytoscape. National resource for network biology [Internet]. 2001. Available from: <http://www.cytoscape.org/>.

Daniels JJ, Autenrieth IB, Ludwig A, Goebel W. The gene slyA of Salmonella typhimurium is required for destruction of M cells and intracellular survival but not for invasion or colonization of the murine small intestine. *Infect Immun.* 1996;64(12):5075-84.

DiLeo MV, Strahan GD, den Bakker M, Hoekenga OA. Weighted correlation network analysis (WGCNA) applied to the tomato fruit metabolome. *PLoS One.* 2011;6(10):e26683.

Domingo-Fernández R, Coll RC, Kearney J, Breit S, O'Neill LAJ. The intracellular chloride channel proteins CLIC1 and CLIC4 induce IL-1 $\beta$  transcription and activate the NLRP3 inflammasome. *J Biol Chem.* 2017;292(29):12077-87.

dos Santos LF, Goncalves EM, Vaz TMI, Guth BEC. Distinct pathotypes of O113 Escherichia coli strains isolated from humans and animals in Brazil. *Journal of Clinical Microbiology.* 2007;45(6):2028-30.

dos Santos LF, Irino K, Vaz TM, Guth BE. Set of virulence genes and genetic relatedness of O113 : H21 Escherichia coli strains isolated from the animal reservoir and human infections in Brazil. *J Med Microbiol.* 2010;59(Pt 6):634-40.

Doughty S, Sloan J, Bennett-Wood V, Robertson M, Robins-Browne RM, Hartland EL. Identification of a novel fimbrial gene cluster related to long polar fimbriae in locus of enterocyte effacement-negative strains of enterohemorrhagic Escherichia coli. *Infection and Immunity.* 2002;70(12):6761-9.

Ellison DW, Miller VL. Regulation of virulence by members of the MarR/SlyA family. *Curr Opin Microbiol*. 2006;9(2):153-9.

Elo LL, Schwikowski B. Analysis of time-resolved gene expression measurements across individuals. *PLoS One*. 2013;8(12):e82340.

Enrichr. Chen E, Kuleshov M, Jones M, Tan C, Kou Y, Clark N, et al [Internet]. Available from: <http://amp.pharm.mssm.edu/Enrichr/>.

Feng P, Delannoy S, Lacher DW, Bosilevac JM, Fach P. Characterization and Virulence Potential of Serogroup O113 Shiga Toxin-Producing *Escherichia coli* Strains Isolated from Beef and Cattle in the United States. *J Food Prot*. 2017;80(3):383-91.

Feng PCH, Delannoy S, Lacher DW, dos Santos LF, Beutin L, Fach P, et al. Genetic Diversity and Virulence Potential of Shiga Toxin-Producing *Escherichia coli* O113:H21 Strains Isolated from Clinical, Environmental, and Food Sources. *Applied and Environmental Microbiology*. 2014;80(15):4757-63.

Feng PCH, Reddy S. Prevalences of Shiga Toxin Subtypes and Selected Other Virulence Factors among Shiga-Toxigenic *Escherichia coli* Strains Isolated from Fresh Produce. *Applied and Environmental Microbiology*. 2013;79(22):6917-23.

Franz E, van Hoek AH, Wuite M, van der Wal FJ, de Boer AG, Bouw EI, et al. Molecular hazard identification of non-O157 Shiga toxin-producing *Escherichia coli* (STEC). *PLoS One*. 2015;10(3):e0120353.

Fratamico PM, DebRoy C, Liu Y, Needleman DS, Baranzoni GM, Feng P. Advances in Molecular Serotyping and Subtyping of *Escherichia coli*. *Front Microbiol*. 2016;7:644.

Fuchs E, Haas AK, Spooner RA, Yoshimura S, Lord JM, Barr FA. Specific Rab GTPase-activating proteins define the Shiga toxin and epidermal growth factor uptake pathways. *J Cell Biol*. 2007;177(6):1133-43.

Funk J, Stoeber H, Hauser E, Schmidt H. Molecular analysis of subtilase cytotoxin genes of food-borne Shiga toxin-producing *Escherichia coli* reveals a new allelic subAB variant. *Bmc Microbiology*. 2013;13.

Gene expression omnibus. National Center for Biotechnology Information, NCBI [internet]. Available from: <http://www.ncbi.nlm.nih.gov/geo>.

Goncalves A, Bürckstümmer T, Dixit E, Scheicher R, Górna MW, Karayel E, et al. Functional dissection of the TBK1 molecular network. *PLoS One*. 2011;6(9):e23971.

Gonzalez AG, Cerqueira AM, Guth BE, Coutinho CA, Liberal MH, Souza RM, et al. Serotypes, virulence markers and cell invasion ability of Shiga toxin-producing *Escherichia coli* strains isolated from healthy dairy cattle. *J Appl Microbiol*. 2016;121(4):1130-43.

Goueli BS, Powell MB, Finger EC, Pfeffer SR. TBC1D16 is a Rab4A GTPase activating protein that regulates receptor recycling and EGF receptor signaling. *Proc Natl Acad Sci U S A*. 2012;109(39):15787-92.

Grizzi F, Franceschini B, Di Biccari S, Musardo S, Pedretti E, Chiriva-Internati M, et al. Sperm protein 17 and AKAP-associated sperm protein cancer/testis antigens are expressed in ciliated hepatic foregut cysts. *Histopathology*. 2015;67(3):398-403.

Guirro M, Fontes Piazza RM, de Souza RL, Cabilio Guth BE. Humoral immune response to Shiga Toxin 2 (Stx2) among Brazilian urban children with hemolytic uremic syndrome and healthy controls. *Bmc Infectious Diseases*. 2014;14.

Gyles CL. Shiga toxin-producing *Escherichia coli*: An overview. *Journal of Animal Science*. 2007;85:E45-E62.

Halbleib JM, Sääf AM, Brown PO, Nelson WJ. Transcriptional modulation of genes encoding structural characteristics of differentiating enterocytes during development of a polarized epithelium in vitro. *Mol Biol Cell*. 2007;18(11):4261-78.

Hartwig JH, Thelen M, Rosen A, Janmey PA, Nairn AC, Aderem A. MARCKS is an actin filament crosslinking protein regulated by protein kinase C and calcium-calmodulin. *Nature*. 1992;356(6370):618-22.

Helgason E, Phung QT, Dueber EC. Recent insights into the complexity of Tank-binding kinase 1 signaling networks: the emerging role of cellular localization in the activation and substrate specificity of TBK1. *FEBS Lett*. 2013;587(8):1230-7.

Henneke P, Golenbock DT. Phagocytosis, innate immunity, and host-pathogen specificity. *J Exp Med*. 2004;199(1):1-4.

Herold S, Paton JC, Paton AW. Sab, a Novel Autotransporter of Locus of Enterocyte Effacement-Negative Shiga-Toxigenic *Escherichia coli* O113:H21, Contributes to Adherence and Biofilm Formation. *Infection and Immunity*. 2009;77(8):3234-43.

Hillmer RA. Systems biology for biologists. *PLoS Pathog*. 2015;11(5):e1004786.

Hopper DW, Ghan R, Schlauch KA, Cramer GR. Transcriptomic network analyses of leaf dehydration responses identify highly connected ABA and

ethylene signaling hubs in three grapevine species differing in drought tolerance. *BMC Plant Biol.* 2016;16(1):118.

Ideker T, Krogan NJ. Differential network biology. *Molecular Systems Biology.* 2012;8.

Irino K, Kato M, Vaz TMI, Ramos, II, Souza MAC, Cruz AS, et al. Serotypes and virulence markers of Shiga toxin-producing *Escherichia coli* (STEC) isolated from dairy cattle in Sao Paulo State, Brazil. *Veterinary Microbiology.* 2005;105(1):29-36.

Ishiwata RR, Morioka MS, Ogishima S, Tanaka H. BioCichlid: central dogma-based 3D visualization system of time-course microarray data on a hierarchical biological network. *Bioinformatics.* 2009;25(4):543-4.

Jandhyala DM, Ahluwalia A, Schimmel JJ, Rogers AB, Leong JM, Thorpe CM. Activation of the Classical Mitogen-Activated Protein Kinases Is Part of the Shiga Toxin-Induced Ribotoxic Stress Response and May Contribute to Shiga Toxin-Induced Inflammation. *Infect Immun.* 2015;84(1):138-48.

Kaper JB, Nataro JP, Mobley HLT. Pathogenic *Escherichia coli*. *Nature Reviews Microbiology.* 2004;2(2):123-40.

Karmali MA, Petric M, Lim C, Fleming PC, Arbus GS, Lior H. The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. *J Infect Dis.* 1985;151(5):775-82.

Kelsall IR, Langenick J, MacKay C, Patel KJ, Alpi AF. The Fanconi anaemia components UBE2T and FANCM are functionally linked to nucleotide excision repair. *PLoS One.* 2012;7(5):e36970.

Krzystek-Korpaczka M, Hotowy K, Czapinska E, Podkowik M, Bania J, Gamian A, et al. Serum availability affects expression of common house-keeping genes in colon adenocarcinoma cell lines: implications for quantitative real-time PCR studies. *Cytotechnology.* 2016;68(6):2503-17.

Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.

Lapointe TK, Buret AG. Interleukin-18 facilitates neutrophil transmigration via myosin light chain kinase-dependent disruption of occludin, without altering epithelial permeability. *Am J Physiol Gastrointest Liver Physiol.* 2012;302(3):G343-51.

Larionov A, Krause A, Miller W. A standard curve based method for relative real time PCR data processing. *BMC Bioinformatics.* 2005;6:62.



- Lechuga S, Naydenov NG, Feygin A, Jimenez AJ, Ivanov AI. A vesicle trafficking protein  $\alpha$ SNAP regulates Paneth cell differentiation in vivo. *Biochem Biophys Res Commun*. 2017;486(4):951-7.
- Lee MS, Cherla RP, Tesh VL. Shiga toxins: intracellular trafficking to the ER leading to activation of host cell stress responses. *Toxins (Basel)*. 2010;2(6):1515-35.
- Lee MS, Kim YJ. Signaling pathways downstream of pattern-recognition receptors and their cross talk. *Annu Rev Biochem*. 2007;76:447-80.
- Lee SM, Suk K, Lee WH. Myristoylated alanine-rich C kinase substrate (MARCKS) regulates the expression of proinflammatory cytokines in macrophages through activation of p38/JNK MAPK and NF- $\kappa$ B. *Cell Immunol*. 2015;296(2):115-21.
- Lesk AM. *Introdução à Bioinformática*. 2nd ed. Ribeirão Preto: Holos; 2008.
- Lin CC, Lee CH, Fuh CS, Juan HF, Huang HC. Link clustering reveals structural characteristics and biological contexts in signed molecular networks. *PLoS One*. 2013;8(6):e67089.
- Liu Q, Jiang C, Xu J, Zhao MT, Van Bortle K, Cheng X, et al. Genome-Wide Temporal Profiling of Transcriptome and Open Chromatin of Early Cardiomyocyte Differentiation Derived From hiPSCs and hESCs. *Circ Res*. 2017;121(4):376-91.
- Liu Y, Yan X, DebRoy C, Fratamico PM, Needleman DS, Li RW, et al. Escherichia coli O-Antigen Gene Clusters of Serogroups O62, O68, O131, O140, O142, and O163: DNA Sequences and Similarity between O62 and O68, and PCR-Based Serogrouping. *Biosensors (Basel)*. 2015;5(1):51-68.
- Liu YY, Slotine JJ, Barabási AL. Controllability of complex networks. *Nature*. 2011;473(7346):167-73.
- Loukil A, Cheung CT, Bendris N, Lemmers B, Peter M, Blanchard JM. Cyclin A2: At the crossroads of cell cycle and cell invasion. *World J Biol Chem*. 2015;6(4):346-50.
- Madsen H. *Time Series Analysis*. Boca Raton, Florida: CRC Press; 2007. p.373.
- Majowicz SE, Scallan E, Jones-Bitton A, Sargeant JM, Stapleton J, Angulo FJ, et al. Global Incidence of Human Shiga Toxin-Producing Escherichia coli Infections and Deaths: A Systematic Review and Knowledge Synthesis. *Foodborne Pathogens and Disease*. 2014;11(6):447-55.

Mao R, Yang R, Chen X, Harhaj EW, Wang X, Fan Y. Regnase-1, a rapid response ribonuclease regulating inflammation and stress responses. *Cell Mol Immunol*. 2017;14(5):412-22.

Matussek A, Einemo IM, Jogenfors A, Löfdahl S, Löfgren S. Shiga Toxin-Producing *Escherichia coli* in Diarrheal Stool of Swedish Children: Evaluation of Polymerase Chain Reaction Screening and Duration of Shiga Toxin Shedding. *J Pediatric Infect Dis Soc*. 2016;5(2):147-51.

May RC, Caron E, Hall A, Machesky LM. Involvement of the Arp2/3 complex in phagocytosis mediated by FcγR or CR3. *Nat Cell Biol*. 2000;2(4):246-8.

Monaghan A, Byrne B, Fanning S, Sweeney T, McDowell D, Bolton DJ. Serotypes and Virulence Profiles of Non-O157 Shiga Toxin-Producing *Escherichia coli* Isolates from Bovine Farms. *Applied and Environmental Microbiology*. 2011;77(24):8662-8.

Monaghan AM, Byrne B, McDowell D, Carroll AM, McNamara EB, Bolton DJ. Characterization of farm, food, and clinical Shiga toxin-producing *Escherichia coli* (STEC) O113. *Foodborne Pathog Dis*. 2012;9(12):1088-96.

Morales SA, Mareninov S, Coulam P, Wadehra M, Goodglick L, Braun J, et al. Functional consequences of interactions between FAK and epithelial membrane protein 2 (EMP2). *Invest Ophthalmol Vis Sci*. 2009;50(10):4949-56.

Moreira-Filho CA, Bando SY, Bertonha FB, Iamashita P, Silva FN, Costa LaF, et al. Community structure analysis of transcriptional networks reveals distinct molecular pathways for early- and late-onset temporal lobe epilepsy with childhood febrile seizures. *PLoS One*. 2015;10(5):e0128174.

Moreira-Filho CA, Bando SY, Bertonha FB, Silva FN, Costa L da F, Carneiro-Sampaio M. Thymus gene coexpression networks: a comparative study in children with and without down syndrome. Passos G. *Transcriptomics in Health and Disease*. Verlag, Berlin, Heidelberg, New York: Springer; 2014a. Cap. 7, p. 123-36.

Moreira-Filho CA, Bando SY, Bertonha FB, Silva FN, Costa L da F. Methods for gene coexpression network visualization and analysis. Passos G. *Transcriptomics in Health and Disease*. Verlag, Berlin, Heidelberg, New York: Springer; 2014b. Cap. 4, p. 79-94.

Moreira-Filho CA, Bando SY, Bertonha FB, Silva FN, Costa LaF, Ferreira LR, et al. Modular transcriptional repertoire and MicroRNA target analyses characterize genomic dysregulation in the thymus of Down syndrome infants. *Oncotarget*. 2016;7(7):7497-533.

Murata Y, Kotani T, Supriatna Y, Kitamura Y, Imada S, Kawahara K, et al. Protein tyrosine phosphatase SAP-1 protects against colitis through regulation of CEACAM20 in the intestinal epithelium. *Proc Natl Acad Sci U S A*. 2015;112(31):E4264-71.

Mysorekar IU, Mulvey MA, Hultgren SJ, Gordon JI. Molecular regulation of urothelial renewal and host defenses during infection with uropathogenic *Escherichia coli*. *J Biol Chem*. 2002;277(9):7412-9.

Naishiro Y, Yamada T, Idogawa M, Honda K, Takada M, Kondo T, et al. Morphological and transcriptional responses of untransformed intestinal epithelial cells to an oncogenic beta-catenin protein. *Oncogene*. 2005;24(19):3141-53.

Nataro JP, Kaper JB. Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev*. 1998;11(1):142-201.

Networks 3D visualization tools for mac [Internet]. Silva FN, Costa L da N. 2016. Available from: <http://cyvision.ifsc.usp.br/software/networks3d>.

Newman MEJ, Girvan M. Finding and evaluating community structure in networks. *Physical Review E*. 2004;69(2).

Newman MEJ. *Networks: an introduction*. New York: Oxford University Press; 2010.

Oshima S, Nakamura T, Namiki S, Okada E, Tsuchiya K, Okamoto R, et al. Interferon regulatory factor 1 (IRF-1) and IRF-2 distinctively up-regulate gene expression and production of interleukin-7 in human intestinal epithelial cells. *Mol Cell Biol*. 2004;24(14):6298-310.

Paton AW, Paton JC. *Escherichia coli* Subtilase Cytotoxin. *Toxins*. 2010;2(2):215-28.

Paton AW, Srimanote P, Talbot UM, Wang H, Paton JC. A new family of potent AB(5) cytotoxins produced by Shiga toxigenic *Escherichia coli*. *Journal of Experimental Medicine*. 2004;200(1):35-46.

Paton AW, Srimanote P, Woodrow MC, Paton JC. Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative shiga-toxigenic *Escherichia coli* strains that are virulent for humans. *Infection and Immunity*. 2001;69(11):6999-7009.

Paton AW, Woodrow MC, Doyle RM, Lanser JA, Paton JC. Molecular characterization of a Shiga toxigenic *Escherichia coli* O113:H21 strain lacking *eae* responsible for a cluster of cases of hemolytic-uremic syndrome. *J Clin Microbiol*. 1999;37(10):3357-61.

- Paton JC, Paton AW. Pathogenesis and diagnosis of shiga toxin-producing *Escherichia coli* infections. *Clinical Microbiology Reviews*. 1998;11(3):450-+.
- Pavlopoulos GA, O'Donoghue SI, Satagopam VP, Soldatos TG, Pafilis E, Schneider R. Arena3D: visualization of biological networks in 3D. *BMC Syst Biol*. 2008;2:104.
- Prifti E, Zucker JD, Clement K, Henegar C. FunNet: an integrative tool for exploring transcriptional interactions. *Bioinformatics*. 2008;24(22):2636-8.
- Primer3Plus. Untergasser A [Internet]. Available from: <http://primer3plus.com/>.
- Qiu W, Wu B, Wang X, Buchanan ME, Regueiro MD, Hartman DJ, et al. PUMA-mediated intestinal epithelial apoptosis contributes to ulcerative colitis in humans and mice. *J Clin Invest*. 2011;121(5):1722-32.
- Rosenkrantz JT, Aarts H, Abee T, Rolfe MD, Knudsen GM, Nielsen M-B, et al. Non-essential genes form the hubs of genome scale protein function and environmental gene expression networks in *Salmonella enterica* serovar Typhimurium. *Bmc Microbiology*. 2013;13.
- Rubino S, Cappuccinelli P, Kelvin DJ. *Escherichia coli* (STEC) serotype O104 outbreak causing haemolytic syndrome (HUS) in Germany and France. *Journal of Infection in Developing Countries*. 2011;5(6):437-40.
- Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, et al. TM4: A free, open-source system for microarray data management and analysis. *Biotechniques*. 2003;34(2):374-+.
- Sahni N, Yi S, Taipale M, Fuxman Bass JI, Coulombe-Huntington J, Yang F, et al. Widespread macromolecular interaction perturbations in human genetic disorders. *Cell*. 2015;161(3):647-60.
- Sahni N, Yi S, Zhong Q, Jaikhani N, Charlotheaux B, Cusick ME, et al. Edgotype: a fundamental link between genotype and phenotype. *Current Opinion in Genetics & Development*. 2013;23(6):649-57.
- Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S, et al. A travel guide to Cytoscape plugins. *Nat Methods*. 2012;9(11):1069-76.
- Sanada T, Takaesu G, Mashima R, Yoshida R, Kobayashi T, Yoshimura A. FLN29 deficiency reveals its negative regulatory role in the Toll-like receptor (TLR) and retinoic acid-inducible gene I (RIG-I)-like helicase signaling pathway. *J Biol Chem*. 2008;283(49):33858-64.

- Scheutz F, Teel LD, Beutin L, Piérard D, Buvens G, Karch H, et al. Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *J Clin Microbiol.* 2012;50(9):2951-63.
- Shimizu M, Kuroda M, Sakashita N, Konishi M, Kaneda H, Igarashi N, et al. Cytokine profiles of patients with enterohemorrhagic *Escherichia coli* O111-induced hemolytic-uremic syndrome. *Cytokine.* 2012;60(3):694-700.
- Sieberts SK, Schadt EE. Moving toward a system genetics view of disease. *Mammalian Genome.* 2007;18(6-7):389-401.
- Silva FN, Amancio DR, Bardosova M, Costa LdaF, Oliveira Jr ON. Using network science and text analytics to produce surveys in a scientific topic. *Journal of Informetrics.* 2016;10(2):487-502.
- Silva FN, Comin CH, Peron TKD, Rodrigues FA, Ye C, Wilson RC et al. Modular Dynamics of Financial Market Networks. *Physics and Society.* 2015;1-13.
- Silva FN, Rodrigues FA, Oliveira Jr ON, Costa LdaF. Quantifying the interdisciplinarity of scientific journals and fields. *Journal of Informetrics.* 2013;7(2):469-77.
- Smyth, G. K. Limma: linear models for microarray data. R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, W. Huber. *Bioinformatics and Computational Biology Solutions using R and Bioconductor*, New York: Springer; 2005. Cap. 23, p. 397-420.
- Song L, Langfelder P, Horvath S. Comparison of co-expression measures: mutual information, correlation, and model based indices. *BMC Bioinformatics.* 2012;13:328.
- Spory A, Bosserhoff A, von Rhein C, Goebel W, Ludwig A. Differential regulation of multiple proteins of *Escherichia coli* and *Salmonella enterica* serovar Typhimurium by the transcriptional regulator SlyA. *J Bacteriol.* 2002;184(13):3549-59.
- Steury MD, Lucas PC, McCabe LR, Parameswaran N. G-protein-coupled receptor kinase-2 is a critical regulator of TNF $\alpha$  signaling in colon epithelial cells. *Biochem J.* 2017;474(14):2301-13.
- Tang B, Li Q, Zhao XH, Wang HG, Li N, Fang Y, et al. Shiga toxins induce autophagic cell death in intestinal epithelial cells via the endoplasmic reticulum stress pathway. *Autophagy.* 2015;11(2):344-54.
- Thorpe CM, Smith WE, Hurley BP, Acheson DW. Shiga toxins induce, superinduce, and stabilize a variety of C-X-C chemokine mRNAs in intestinal

epithelial cells, resulting in increased chemokine expression. *Infect Immun.* 2001;69(10):6140-7.

Tran HJ, Heroven AK, Winkler L, Spreter T, Beatrix B, Dersch P. Analysis of RovA, a transcriptional regulator of *Yersinia pseudotuberculosis* virulence that acts through antirepression and direct transcriptional activation. *J Biol Chem.* 2005;280(51):42423-32.

Vareille M, Rannou F, Théliier N, Glasser AL, de Sablet T, Martin C, et al. Heme oxygenase-1 is a critical regulator of nitric oxide production in enterohemorrhagic *Escherichia coli*-infected human enterocytes. *J Immunol.* 2008;180(8):5720-6.

Vaz TMI, Irino K, Kato M, Dias AMG, Gomes TAT, Medeiros MIC, et al. Virulence properties and characteristics of Shiga toxin-producing *Escherichia coli* in Sao Paulo, Brazil, from 1976 through 1999. *Journal of Clinical Microbiology.* 2004;42(2):903-+.

Wall EA, Zavzavadjian JR, Chang MS, Randhawa B, Zhu X, Hsueh RC, et al. Suppression of LPS-induced TNF-alpha production in macrophages by cAMP is mediated by PKA-AKAP95-p105. *Sci Signal.* 2009;2(75):ra28.

Wang LS, Xiao XY, Li DL, Chi YY, Wei P, Wang YQ, et al. Abnormal expression of GADD45B in human colorectal carcinoma. *Journal of Translational Medicine.* 2012;10.

Wang X, Kim Y, Ma Q, Hong SH, Pokusaeva K, Sturino JM, et al. Cryptic prophages help bacteria cope with adverse environments. *Nat Commun.* 2010;1:147.

Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE, et al. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol Biol Cell.* 2002;13(6):1977-2000.

Wiedemann A, Linder S, Grassl G, Albert M, Autenrieth I, Aepfelbacher M. *Yersinia enterocolitica* invasin triggers phagocytosis via beta1 integrins, CDC42Hs and WASp in macrophages. *Cell Microbiol.* 2001;3(10):693-702.

Winchell CG, Steele S, Kawula T, Voth DE. Dining in: intracellular bacterial pathogen interplay with autophagy. *Curr Opin Microbiol.* 2016;29:9-14.

Winterbach W, Van Mieghem P, Reinders M, Wang H, de Ridder D. Topology of molecular interaction networks. *Bmc Systems Biology.* 2013;7.

World Health Organization. *World health statistics 2012*, Part.III Global health indicators. Geneva: WHO; 2012.

World Health Organization. *World health statistics 2016*, SDG health and health-related targets. Geneva: WHO; 2016.

Xu Y, Zhang M, Wang Y, Kadambi P, Dave V, Lu LJ, et al. A systems approach to mapping transcriptional networks controlling surfactant homeostasis. *Bmc Genomics*. 2010;11.

Yu T, Shan TD, Li JY, Huang CZ, Wang SY, Ouyang H, et al. Knockdown of linc-UFC1 suppresses proliferation and induces apoptosis of colorectal cancer. *Cell Death Dis*. 2016;7:e2228.

Yu Z, Liu J, Deng WM, Jiao R. Histone chaperone CAF-1: essential roles in multi-cellular organism development. *Cell Mol Life Sci*. 2015;72(2):327-37.

Yun SH, Ji SC, Jeon HJ, Wang X, Kim SW, Bak G, et al. The CnuK9E H-NS Complex Antagonizes DNA Binding of DicA and Leads to Temperature-Dependent Filamentous Growth in E-coli. *Plos One*. 2012;7(9).

Zayat V, Balcerak A, Korczynski J, Trebinska A, Wysocki J, Sarnowska E, et al. HAX-1: a novel p-body protein. *DNA Cell Biol*. 2015;34(1):43-54.

Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol*. 2005;4:Article17.

Zhu H, Wan X, Li J, Han L, Bo X, Chen W, et al. Computational Prediction and Validation of BAHD1 as a Novel Molecule for Ulcerative Colitis. *Sci Rep*. 2015;5:12227.

Zhu X, Gerstein M, Snyder M. Getting connected: analysis and principles of biological networks. *Genes & Development*. 2007;21(9):1010-24.

Zundler S, Caioni M, Müller M, Strauch U, Kunst C, Woelfel G. K<sup>+</sup> Channel Inhibition Differentially Regulates Migration of Intestinal Epithelial Cells in Inflamed vs. Non-Inflamed Conditions in a PI3K/Akt-Mediated Manner. *PLoS One*. 2016;11(1):e0147736.