

Article

Enteroinvasive *Escherichia Coli* is Inefficient to Escape from Autophagy during Caco-2 Cellular Infection

Silvia Yumi Bando¹, Ana Carolina Ramos Moreno², Hadassa Cristhina de Azevedo Soares dos Santos³, Fernanda Bernardi Bertonha¹, Lucas Gonçalves Ferreira³, Marina Baquerizo Martinez³, and Carlos Alberto Moreira-Filho^{1*}

¹ Department of Pediatrics, Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil; e-mail: silvia.bando@fm.usp.br (SYB); e-mail: fernanda.bernardi@fm.usp.br (FBB); e-mail: carlos.moreira@fm.usp.br (CAMF)

² Department of Microbiology, Institute of Biomedical Sciences, Universidade de São Paulo, São Paulo, SP, Brazil; e-mail: carol@usp.br (ACRM)

³ Department of Clinical and Toxicological Analyses, Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, São Paulo, SP, Brazil; e-mail: hadassa.cass@gmail.com (HCASS); e-mail: lucasgferreira@yahoo.com.br (LGF); e-mail: mabanez74@gmail.com (MBM)

* Correspondence: e-mail: carlos.moreira@hc.fm.usp.br

Abstract: *Escherichia coli* and *Shigella* are common diarrhea-causing pathogens in children and adults. Enteroinvasive *Escherichia coli* (EIEC) shares a similar pathogenic mechanism with *Shigella*. However, EIEC are less virulent than *Shigella*. The aim of this work was to get a better understanding of the virulence differences between EIEC and *S. flexneri*. We investigated i) the bacterial gene co-expression networks (GCNs) and ii) the transcriptional modules (WGCNA) of Caco-2 cells infected with EIEC or with *S. flexneri* during a three-hour period of bacterial infection. The GCN analysis showed that EIEC and *S. flexneri* networks presented different topologies. Additionally, the EIEC network revealed that pINV genes are not connected with chromosomal genes. WGCNA and eigengene analysis showed enterocyte gene expression variation along the three-hour bacterial post-infection period. Additionally, at one-hour post-infection EIEC induced a higher number of gene expression changes in Caco-2 cells than *S. flexneri*. Several of these genes are involved in autophagy. This study showed that the lower virulence of EIEC is associated with a lack of functional cooperation between pINV and chromosomal genes, differently from what was observed in *S. flexneri*. Consequently, EIEC becomes less efficient in subverting host-cell bacterial recognition as well as defense mechanisms such as autophagy.

Keywords: Enteroinvasive *Escherichia coli*; *Shigella flexneri*; Caco-2 cells; icsB; autophagy; WGCNA; gene co-expression network

1. Introduction

Infectious diarrhea is one of the main causes of childhood death, especially in developing countries [1] and *Escherichia coli* and *Shigella* are common diarrhea-causing pathogens in children and adults [2]. Enteroinvasive *Escherichia coli* (EIEC) is an intestinal pathogen causing enteritis [3] and shares a similar pathogenic mechanism with *Shigella*, characterized by the destruction of the colonic epithelium triggered by mucosal inflammation and ulceration upon bacterial invasion [4,5]. However, EIEC strains are less virulent than *Shigella* strains, since a higher infectious dose of EIEC is needed to cause disease in volunteers [6]. This fact could be in part explained by EIEC lower expression of virulence genes and by distinct host recognition mechanisms [7-11]. Moreover, it was observed that EIEC has diminished intestinal cell-to-cell spreading capability compared to *S. flexneri* [11].

The virulence genes required for bacterial-cell infection are located in the virulence plasmid pINV present in *Shigella* and EIEC [12-14]. This plasmid harbor genes codifying proteins that are crucial for bacteria entrance into host cells (*ipa* genes) and intracellular

survival and spread (*ics* and *vir* genes). The protein IcsA/VirG (intra cellular spread) is one of the mediators of this process [15-17]. Yet, another bacterial protein, called IcsB, is required, which binds to a region of the IcsA, preventing its recognition by the autophagy protein Atg5 and, consequently, avoiding pathogen degradation by xenophagy [18, 19].

Transcriptomic and gene co-expression network analyses have been successfully used for unveiling the molecular mechanisms involved in pathogenic *E. coli* virulence [20, 21]. Thus, to get a better understanding of the virulence differences between EIEC and *S. flexneri* we investigated the transcriptomic mechanisms involved in the responses of Caco-2 cells infected with EIEC or with *S. flexneri*. Two experimental approaches were adopted to fulfill this objective: i) comparative gene co-expression network analysis of EIEC and *S. flexneri* bacteria exposed to fresh media or to Caco-2 conditioned media; ii) transcriptional module analysis of Caco-2 cells responses to EIEC or *S. flexneri* during a three-hour period of bacterial infection.

2. Materials and Methods

2.1. Bacterial strains

EIEC O124:H- (FBC124/13) and *S. flexneri* serotype 5a (M90T) strains are kept at the Department of Clinical Chemistry and Toxicology, Faculdade de Ciências Farmacêuticas, USP, São Paulo, SP, Brazil. EIEC FBC124/13 was isolated from a child with diarrhea [9], and M90T is a wild-type strain of *S. flexneri*, provided by Philippe Sansonetti, Institut Pasteur, Paris, France. The serotype, virulence factors, and Serény test were previously described [9]. Both strains presented positive Serény and harbor pINV [9]. For all bacterium-enterocyte interaction assays described below, strains were grown in Congo red agar plates and a single red colony was cultured with shaking for 4 hours in 3 mL of tryptic soy broth medium (Difco) at 37° C. Subsequently, the bacterial cultures were spun and 1x10⁸ CFU/mL (0.1 OD 625nm) were then suspended in DMEM (Sigma-Aldrich).

2.2. Caco-2 cell culture conditions, bacterial infection assays, and RNA extraction

We used Caco-2 cells (ATCC® HTB-37™), a human epithelial cell line from colon tissue that mimics the intestinal epithelium in vitro [22]. The Caco-2 cells were cultured in six-well tissue culture plates containing Dulbecco's Modified Eagle's Medium (DMEM) with 10% fetal bovine serum (FBS, Invitrogen, Carlsbad, CA), 1mM L-glutamine (Sigma-Aldrich), penicillin-streptomycin (100 U/mL-100 mg/mL, Sigma-Aldrich), and 1% non-essential amino acids. The culture plates were incubated in a 5% CO₂ at 37° C. The cells were grown until confluence and the formation of a polarized epithelial cell monolayer (after five to seven days in culture). Prior to bacterial infection assays, cells were gently washed with phosphate-buffered saline (PBS). Subsequently, 1 mL of antibiotic-free DMEM with 10% FBS and 100 µL of bacterial culture (1x10⁸CFU/mL) was added to each well, centrifuged for 5 min at 100 x g to bacteria-cell contact, and incubated at 37° C in 5% CO₂ for two hours, the time required bacterial invasion [11]. After this period, the cells were washed once with 1 mL of PBS, fresh medium supplemented with 50 µg/mL of gentamycin was added and cell cultures were incubated in 5% CO₂ at 37°C for 30 min to kill extracellular bacteria. After washing with PBS, 1 ml of fresh DMEM with 10% FBS was added, and the cultures were incubated at 37° C in 5% CO₂ for different time intervals. For the bacteria-free control group the culturing and washing procedures were the same but performed with antibiotic-free DMEM with 10% FBS. After each time interval, Caco-2 cells were recovered for RNA extraction (see below). The time series encompassed three sequential intervals of one hour each (T0 to T2). In T0 the Caco-2 cells were recovered immediately after removing the gentamycin-DMEM medium. All bacterial infection assays and control group were done in quadruplicates.

After each incubation period, the cultured Caco-2 cells were washed with PBS and, subsequently, directly lysed in the culture wells with 700 µL of Buffer RLT (Qiagen, Valencia, CA). The cell lysate was collected, and total RNA extraction was accomplished

using the RNeasy Mini kit (Qiagen, Valencia, CA). RNA purity analysis and quantification were done using the NanoVue spectrophotometer (GE Healthcare Life Sciences, Marlborough, MA). RNA quality was assessed on the Agilent BioAnalyzer 2100 (Agilent, Santa Clara, CA). All samples presenting RIN ≥ 9.0 and were stored at -80°C until used in hybridization experiments

2.3. Growth of bacterial strains in fresh or conditioned medium and RNA extraction

The bacterial strains were grown in two different media: i) fresh medium, i.e., antibiotic-free DMEM medium containing 10% fetal bovine serum (FBS); and ii) the same medium recovered after 24 h of incubation with Caco-2 cells, designated conditioned medium. Fresh or conditioned media were subsequently used for bacterial growth assays involving the strains FBC124/13 or M90T. Briefly, 400 μL of bacterial culture (1×10^8 CFU/mL) were added in 4 mL of fresh or conditioned medium and incubated for 3 h at 37°C . There were six biological replicates for each bacterial strain. After this period the bacteria were recovered by centrifugation for 10 min at $5,000 \times g$ and the pellet was suspended in 600 μL of RNeasy Protect Bacteria Reagent (Qiagen cat. no. 76506, Valencia, CA) for RNA preservation. Bacterial cells were lysed using lysozyme (1 mg/mL) and proteinase K (2 mg/mL) and incubated for 10 min at 20°C . Total RNA was obtained using the RNeasy Mini Kit (Qiagen cat no. 74104, Valencia, CA). RNA purity analysis and quantification were accomplished by using the NanoVue spectrophotometer. RNA integrity was assessed on the Agilent BioAnalyzer 2100 (Agilent, Santa Clara, CA). All samples presented RIN > 7.0 and were stored at -80°C until use in hybridization experiments.

2.1. Microarray hybridization

2.4.1. Gene expression profiles for bacterial strains

To determine the gene expression profiles for FBC124/13 and M90T strains, 15K DNA microarrays (customized *E. coli* and pINV genes of *S. flexneri* M90T Gene Expression Microarray, G2509F-046693, Agilent Technologies, Santa Clara, CA) were used. Cyanine 3-CTP fluorescent dye (Cy-3 dye) was used for hybridization labeling (Fairplay III Labeling, version april/2009, Stratagene, adapted for one-color protocol).

2.4.2. Gene expression profiles for bacterial strains

Gene expression profiles for Caco-2 cells were obtained using 4x44K v.2 DNA microarrays (Whole Human Genome Microarray Kit, Agilent Technologies, cat no. G4845A, Santa Clara, CA). Cy-3 dye was used for hybridization labelling (One-Color Microarray-Based Gene Expression Analysis - Low Input Quick Amp Labeling).

2.5. Data acquisition and processing

The images were captured by the reader Agilent Bundle, according to the parameters recommended for bio-arrays and extracted by Agilent Feature Extraction software version 9.5.3. Spots with two or more flags (low intensity, saturation, controls, etc.) were considered as NA, that is, without valid expression value. An in-house algorithm in R environment [23] (version 3.6.3) was used for excluding transcript spots presenting one or more NAs and for converting gene expression values to log base 2. Data normalization was performed using limma package [24] in R environment [23] (version 3.6.3). All microarray raw data were deposited in GEO public database (www.ncbi.nlm.nih.gov/geo), a MIAME compliant database, under accession numbers: GSE93588, GSE93525, and GSE93587 for Caco-2 infected with EIEC FBC124/13, Caco-2 infected with *S. flexneri* M90T, and uninfected Caco-2, respectively.

2.6. Bacterial gene expression data analyses

2.6.1. Differential gene expression analysis

The normalized gene expression data of all transcripts were used for obtaining the differentially expressed genes/transcripts (DE genes). TMEV software version 4.9.0 and Significance Analysis of Microarrays (SAM) was used for all comparisons. The genes or transcripts presenting fold change of 2.0 or higher were considered differentially expressed (DE).

2.6.2. Bacterial gene co-expression networks

Gene co-expression networks (GCNs) were constructed by using Pearson's correlation. Data analysis and visualization were accomplished through Cytoscape [25] (version 3.7.2). All networks were tested for scale-free status, i.e., power-law distributions in empirical data [26]. The GCN correlation thresholds were chosen in order to ensure that most of the nodes remained connected to the major network components and that the network remained stable along a threshold range, i.e., maintaining network's topological structure [27].

2.7. *Caco-2* gene co-expression network analysis

Caco-2 cells temporal gene co-expression variation along three-hour bacterial infection was performed by transcriptional module analysis associated with three equal one-hour intervals. The time series data were grouped in three transcriptomic sets: i) Caco-2 infected with EIEC FBC124/13 (group CE), Caco-2 infected with *S. flexneri* M90T (group CS), and uninfected Caco-2 (group CC). Data normalization was performed using limma package [24] in R environment [23] (version 3.6.3).

2.7.1. Weighted gene co-expression network analysis (WGCNA)

The normalized time series data of Caco-2 cells gene expression were used for WGCNA. Networks for each group were constructed using the WGCNA package [28] in R environment [23] (version 3.4.4). Pearson's correlation coefficient was used for obtaining gene co-expression similarity measures and for the subsequent construction of an adjacency matrix using soft power and topological overlap matrix (TOM). Soft-thresholding process transforms the correlation matrix to mimic the scale-free topology. TOM is used to filter weak connections during network construction. Module identification is based on TOM and in average linkage hierarchical clustering. Keeping to the scale-free topology criterion, soft power $\beta = 20, 16$, and 22 were considered for CE, CS, and CC networks, respectively (Fig. S1-S3). Finally, Dynamic Tree Cut algorithm was used for dendrogram's branch selection. The module eigengene is defined as the first principal component of a given module, which can be considered a representative of the gene expression profiles in a module.

2.7.2. Module-trait association

Each time interval was considered as a trait. Consequently, we obtained the Gene Significance (GS), which is a value of the correlation between the gene expression values and the trait group [28]. Subsequently, the module significance (MS) value is the correlation of the GS values of the module genes and the trait. The modules presenting significantly and high MS value ($MS > 0.85$ and $p < 5e-04$) to time-interval were selected for biological functional analysis of module genes.

2.8. Enrichment analysis

KEGG pathways enrichment analyses on gene sets of the modules highly correlated to the time-intervals was performed using Enrichr online web tool [29, 30].

2.9. Venn diagram analysis

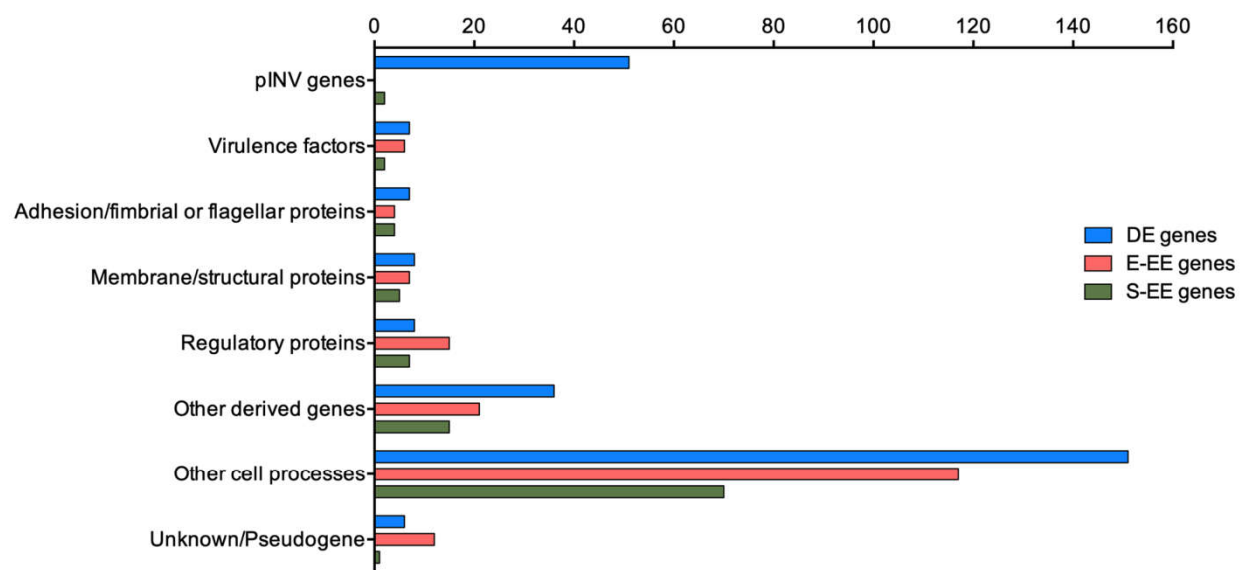
The comparative analyses for all datasets - module genes, or KEGG terms - obtained for CE, CS, and CC groups were performed by online web tool Calculate and Draw Custom Venn Diagrams [31].

3. Results

3.1. Bacterial gene expression analyses

A total of 4,573 genes/transcripts were analyzed. The first analysis aimed at investigating if Caco-2 cells soluble mediators modify bacterial global gene expression. Thus, we performed a comparative analysis by growing the EIEC FBC124/13 and *S. flexneri* M90T strains in conditioned (C) or fresh (F) medium for three hours. Comparative gene expression analysis of these strains cultured in C and F media showed only one gene, *sodA*, that codifies for a superoxide dismutase was DE in *S. flexneri*, while in EIEC two transcripts were DE: ECs3163, a putative outer membrane protein, and *ygaW*, a hypothetical protein. These results showed that EIEC and *S. flexneri* strains presented almost no differences in their transcriptomic profiles after exposure to Caco-2 cells soluble mediators.

Consequently, we conducted a comparison between EIEC FBC124/13 and *S. flexneri* M90T bacterial strains cultured in C medium. The comparative analysis revealed that 273 genes were DE, of which 152 were hypo-expressed in EIEC. Additionally, 183 genes were exclusively expressed in the EIEC strain (E-EE genes), and 104 were exclusively expressed in the *S. flexneri* strain (S-EE genes). Biological function analysis (Fig. 1) showed that several DE genes are pINV-derived genes (51 genes) or foreign/horizontal acquired genes (36 genes), such as prophage, plasmid, insertion element, etc. Among the exclusively expressed genes found in EIEC or *S. flexneri* (Fig. 1), many are foreign/horizontal-acquired genes (21 in EIEC and 15 in *S. flexneri*) or codify for transcriptional regulators (15 in EIEC and 7 in *S. flexneri*). A detailed description of the biological functions related to the DE and to the exclusively expressed genes in the EIEC or *S. flexneri* strains are listed in Table S1.



Functional categories

pINV genes

Virulence factors: Antibiotic resistance protein/Drug binding; Autolysis; Cytotoxic protein; Drug resistance; Macrophage toxin; Type III secretion

Adhesion/fimbrial or flagellar proteins: Adhesion protein; Fimbrial protein; Flagellar protein

Membrane/structural proteins: LPS; Membrane protein; Outer membrane; Structural protein

Regulatory proteins

Other derived genes: Insertion element; Lambda DNA; pO157 gene; Prophage gene; Tail protein; Transposase

Other cell processes: Cell division; DNA replication; Lipoprotein; Metabolic process; pH stress; Protein binding; Transcription; Transport

Unknown/Pseudogene

Figure 1. Distribution of the bacterial genes by functional categories. DE, differentially expressed; genes expressed only in EIEC (E-EE) or in *S. flexneri* (S-EE).

3.1.1. Bacterial gene co-expression network analysis

The GCNs for EIEC FBC124/13 and for *S. flexneri* M90T strains were constructed including the DE genes, the non-DE pINV genes, and the exclusively expressed genes in each strain (Table 1, Table S2). The GCNs obtained for EIEC and *S. flexneri* revealed different topologies for both bacteria. The EIEC has three networks (Fig. 2) whereas the *S. flexneri* has one network (Fig. 3).

Table 1. Number of genes used for EIEC and *S. flexneri* gene co-expression networks construction.

Group	DE		Expressed only in		
	genes	pINV	genes	pINV	NDE-pINV
EIEC	222	51	183	0	20
M90T	222	51	104	2	20

NDE-pINV, pINV genes did not differentially expressed.

The three EIEC networks were named A, B, and C (Fig. 2), as follows. The network A contains the larger number of genes, being essentially constituted by E-EE and DE genes (most of the DE genes are hypo-expressed in EIEC when compared with *S. flexneri*). This network encompasses only six pINV-derived genes. Conversely, the other two small

networks, B and C, contain, respectively, 31% (5 out of 16 genes) and 50% (6 out of 12 genes) of the pINV-derived genes (Table 2). Four of these pINV-derived genes - *ipaA*, *ipaC*, *ipaD*, and *icsB* - are well-known virulence factors. However, the B and C networks are not connected with the major network A. Thirteen genes were identified as hubs (i.e. highly connected genes) and all of them are in the network A (Table 3). Eight of these hubs are still putative genes and the other five are involved in cellular processes or in organelle structures, i.e., metabolism (*yjiL* and *yiaK*), transmembrane transport (*yibH*), cytoskeleton organization (*cbeA*), and ribosomal structure (*rluD*).

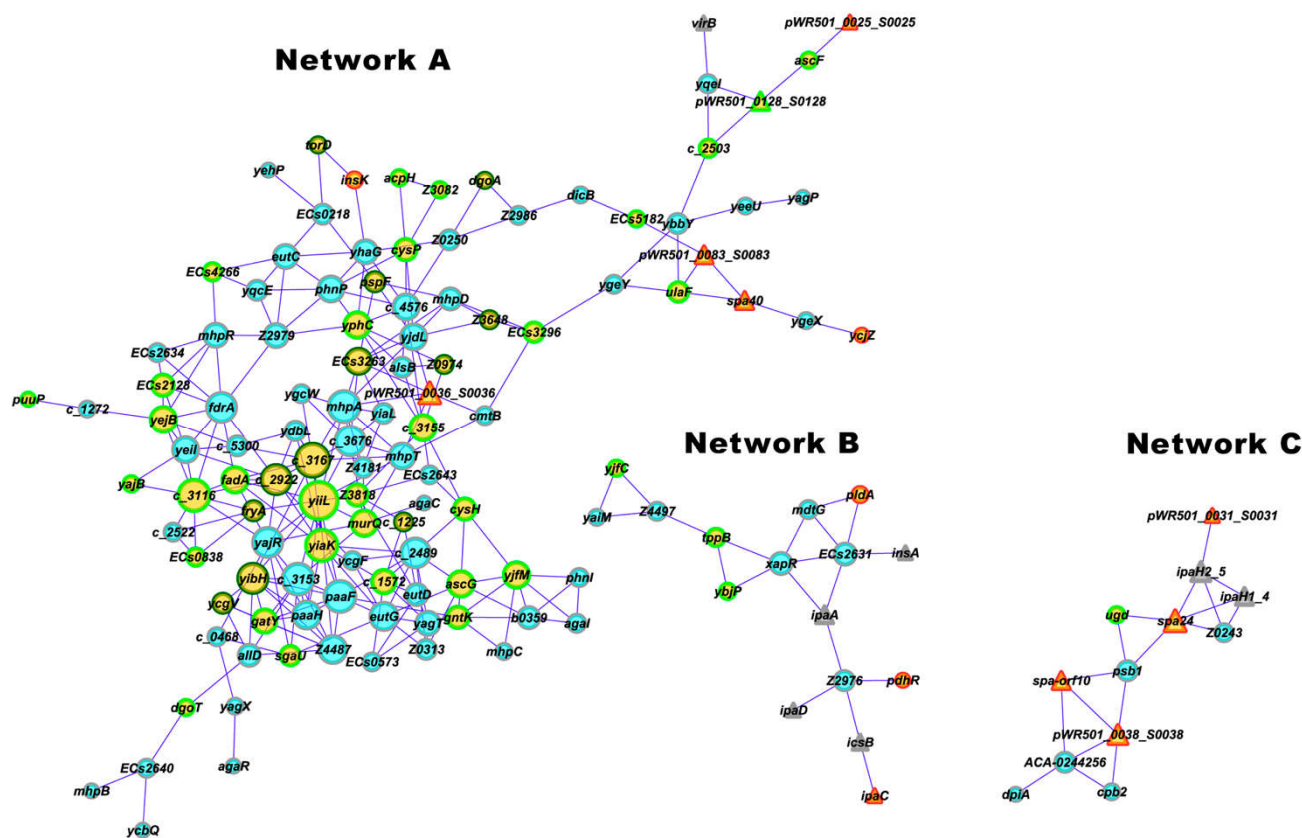


Figure 2. EIEC gene-co-expression networks (A, B, and C). Circle or triangle nodes indicate, respectively, chromosome or pINV genes. Nodes filled with yellow, cyan, or gray colors indicate, respectively, the differentially (DE) genes, exclusive expressed genes or non-DE genes. The hyper-, hypo-, or very hypo-expressed genes are identified by nodes with red, dark green, or green border respectively. The node size is proportional to the number or gene-gene links. All gene-gene links presented positive correlation values.

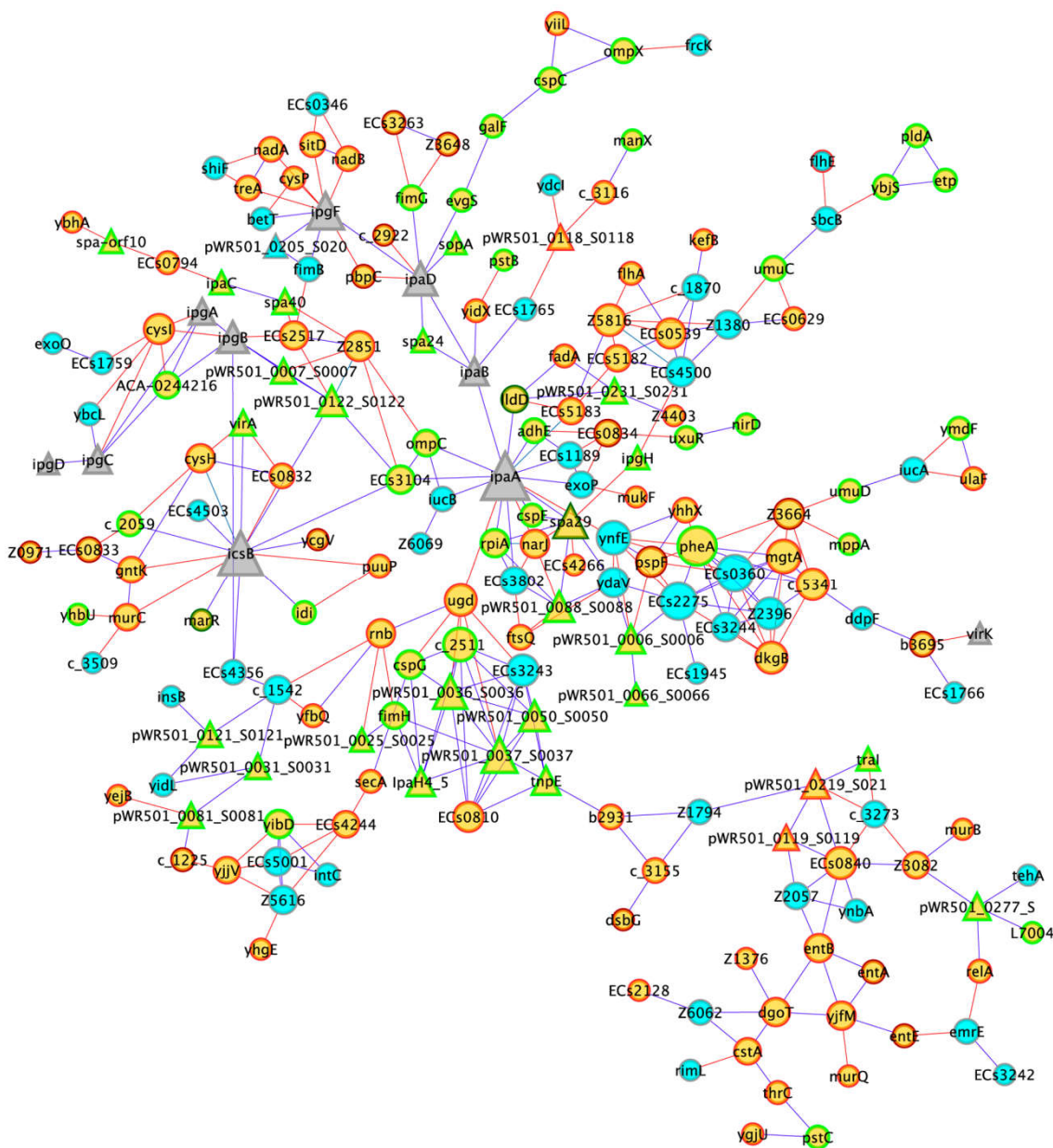


Figure 3. *S. flexneri* gene co-expression network. Circle or triangle nodes indicate, respectively, chromosome or pINV genes. Nodes filled with yellow, cyan, or gray colors indicate, respectively, the differentially (DE) genes, exclusive expressed genes or non-DE genes. The highly hyper-, hyper, very hypo-, or hypo-expressed genes are identified by nodes with dark red, red, dark green, or green border respectively. The node size is proportional to the number of gene-gene links. Gene-gene links presented positive or negative correlation values are indicated in blue or red lines.

The *S. flexneri* network contains several DE genes (hyper-expressed genes in comparison with EIEC) and 20% (39 out of 197 genes) of pINV genes (Fig. 3 and Table 2). In this network seven out of 13 hubs are pINV genes. Among the pINV hubs, three are involved in bacterial invasion (*ipaA*, *IpaD*, and *spa29*), two involved in cell-to-cell spread (*icsB* and *ipgF*), and the other two have yet undescribed functions (Table 4). Other five hubs are involved in cellular process, such as metabolism (*pheA* and *yrfE*), ion transport (*betT* and *mgtA*), and DNA replication (*ECs2275*). Only one is a putative ATP-binding protein (*c_2511*).

Table 2. Gene co-expression network characteristics for EIEC and *S. flexneri* strains.

Group	Cut-off (r ≥)	Network	No. nodes	No. Links	no. of DE genes				E-EE genes	no. of pINV-derived genes				
					Hypo EIEC	in	Hyper EIEC	in		Hypo EIEC	in	Hyper EIEC	in	NDE
EIEC	0.970	A	110	252	42		2		60	1		4		1
		B	16	20	3		2		6	0		1		4
		C	12	18	1		0		5	0		4		2
					Hypo M90T	in	Hyper M90T	in	S-EE genes	Hypo M90T	in	Hyper M90T	in	NDE
<i>S. flexneri</i>	0.965		197	342	34		76		48	25		3		10

DE, differentially expressed genes; NDE, genes were not differentially expressed between EIEC and *S. flexneri*; E-EE or S-EE, genes exclusively expressed in EIEC or *S. flexneri* respectively.

Table 3. Hub genes in the EIEC network A.

Gene	No. Links	Profile	FC	Gene description	Function/structure
<i>yjiL</i>	13	DE	0.134	alias <i>rhaM</i> ; L-rhamnose mutarotase; rhamnose catabolic process	Metabolic process
<i>c_3167</i>	11	DE	0.019	Putative DNA packaging protein of prophage	Prophage
<i>yiaK</i>	10	DE	0.083	2,3-diketo-L-gulonate reductase; oxidation-reduction process	Metabolic process
<i>c_3116</i>	10	DE	0.083	Ribosomal large subunit pseudouridine synthase D	Translation/ribosomal structure
<i>paaF</i>	10	EE	NA	putative 2,3-dehydrodipyl-CoA hydratase; fatty acid beta-oxidation	Metabolic process
<i>c_3153</i>	10	EE	NA	Putative outer membrane protein of prophage	Prophage
<i>mhpA</i>	10	EE	NA	putative 3-(3-hydroxyphenyl)propanoate/3-hydroxycinnamate hydroxylase; 3-(3-hydroxy)phenylpropionate catabolic process	Metabolic process
<i>c_2489</i>	9	EE	NA	<i>yhiX</i> ; putative pyruvate transporter; transmembrane transport	transport
<i>c_2922</i>	9	DE	0.010	<i>fryA</i> ; putative PTS multiphosphoryl transfer protein Fry; carbohydrate transport	Metabolic process
<i>c_3676</i>	9	EE	NA	alias <i>cbeA</i> ; CP4-44 prophage; cytoskeleton bundling-enhancing antitoxin CbeA; positive regulation of cytoskeleton organization	cytoskeleton
<i>yajR</i>	9	EE	NA	putative major facilitator superfamily transport protein YajR; ransmembrane transport	transport
<i>yibH</i>	9	DE	0.019	inner membrane protein YibH; transmembrane transport	transport
<i>fdrA</i>	9	EE	NA	putative acyl-CoA synthetase FdrA	bacterial-type flagellum-dependent cell motility

FC, fold change (EIEC/M90T); DE, differentially expressed genes; EE, exclusively expressed genes.

Table 4. Hub genes in the *S. flexneri* network. The pINV genes are highlighted in grey.

Gene	No. Links	Profile	FC	Gene description	Function/structure
------	--------------	---------	----	------------------	--------------------

<i>ipaA</i>	16	NDE	NA	pWR501 gene; bacterial invasion protein	bacterial invasion
<i>icsB</i>	14	NDE	NA	pWR501 gene; protein secreted by the type III secretion apparatus of <i>Shigella flexneri</i>	cell-to-cell spread
<i>pheA</i>	11	DE	0.169	L-phenylalanine biosynthetic process	metabolic process
<i>ECs0360</i>	10	SE	NA	choline transport protein BetT	ion transport
<i>ECs2275</i>	10	SE	NA	acts to load the DnaB helicase onto the initiation site during DNA replication	DNA replication
<i>ipgF</i>	10	NDE	NA	pWR501 gene; periplasmic protein, similarities to the catalytic site of lysozymes	cell-to-cell spread
pWR501_0037_S0037	9	DE	0.086	pWR501 gene	unknown
c_2511	8	DE	0.094	insertion sequence	putative ATP-binding protein
<i>ynfE</i>	8	SE	NA	putative oxidoreductase	metabolic process
<i>ipaD</i>	8	NDE	NA	pWR501 gene; secreted by the Mxi-Spa machinery, required for entry of bacteria into epithelial cells	bacterial invasion
<i>mgtA</i>	8	DE	8.432	Mg2+ transport ATPase	ion transport
pWR501_0036_S0036	8	DE	0.103	pWR501 gene	unknown
<i>spa29</i>	8	DE	0.022	pWR501 gene; component of the Mxi-Spa secretion machinery	bacterial invasion

FC, fold change (M90T/EIEC); DE, differentially expressed genes; EE, exclusively expressed genes.

3.2. Transcriptional analyses of *Caco-2* cells after bacterial challenge

3.2.1. Network and module-trait correlation analyses

A total of 4,812 GO annotated genes were used for constructing three WGCNA gene networks: one for *Caco-2* infected with EIEC FBC124/13 (CE group), other for *Caco-2* infected with *S. flexneri* M90T (CS group), and the third for *Caco-2* without bacteria (CC group). Hierarchical clustering analyses identified respectively, 11, 12, or 9 distinct gene modules for each group. Module size ranged from 124 (green-yellow module) to 1,129 (turquoise module) genes for the CE group, from 98 (tan module) to 1,561 (turquoise module) genes for the CS group, and from 119 (pink module) to 780 (turquoise module) for the CC group (Figures S1-S3).

The module-trait correlation analysis for the CE group (Fig. S1) revealed that seven modules were significantly ($p < 0.05$) and positively correlated to T0 (turquoise, purple, and yellow modules), T1 (blue and brown modules), and T2 (magenta and pink modules). The analysis for the CS group (Fig. S2) revealed that six modules were significantly ($p < 0.05$) and positively correlated to T0 (turquoise and yellow modules), T1 (magenta module), and T2 (purple, blue, and pink modules). The analysis for the CC group showed that four modules were correlated with two time intervals (Fig. S3): T0 (turquoise and yellow modules) and T2 (blue and red modules).

The module relationship analysis based on the eigengene dendrogram (Fig. 4) showed that for the CE and CS networks there are three metamodules presenting a positive correlation with the T0, T1, and T2 intervals. The CC network one metamodule was correlated to the T0 and T1 intervals and the dissimilar one was correlated to T2 interval. It is interesting to note that for CE and CS networks the metamodule correlated with the T1 interval has a distant relationship with the metamodules correlated with T0 and T2 intervals.

3.1.1. Bacterial gene co-expression network analysis

The GCNs for EIEC FBC124/13 and for *S. flexneri* M90T strains were constructed including the DE genes, the non-DE pINV genes, and the exclusively expressed genes in each strain (Table 1, Table S2). The GCNs obtained for EIEC and *S. flexneri* revealed different topologies for both bacteria. The EIEC has three networks (Fig. 2) whereas the *S. flexneri* has one network (Fig. 3).

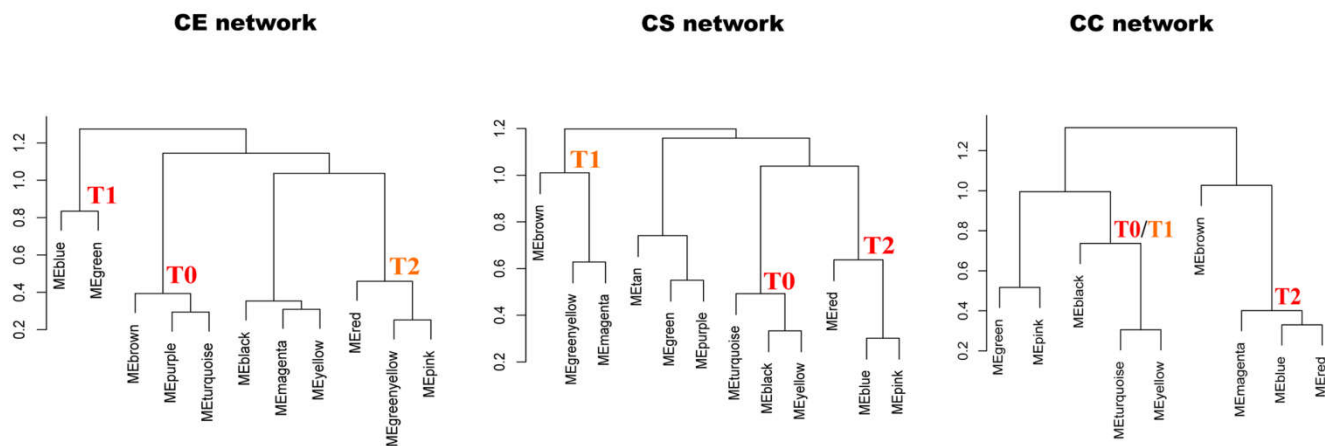


Figure 4. Eigengene analysis. Module-trait relationship for CE, CS, and CC groups. Eigengene dendrograms for the CE, CS, and CC groups. T0, T1, and T2 indicate metamodules positively and highly (red) or positively (orange) correlated with time-intervals.

3.2.2. Functional analysis

We selected the highly ($r > 0.85$) correlated modules with the T0, T1 or T2 intervals in the CE, CS, and CC networks (Fig. 5). Firstly, we compared the gene composition of these modules per time-intervals. This analysis showed that several genes were in the overlapping regions of the diagrams, as particularly seen in the T0 interval. However, in the intervals that follow, a smaller number of genes appear in the intersections.

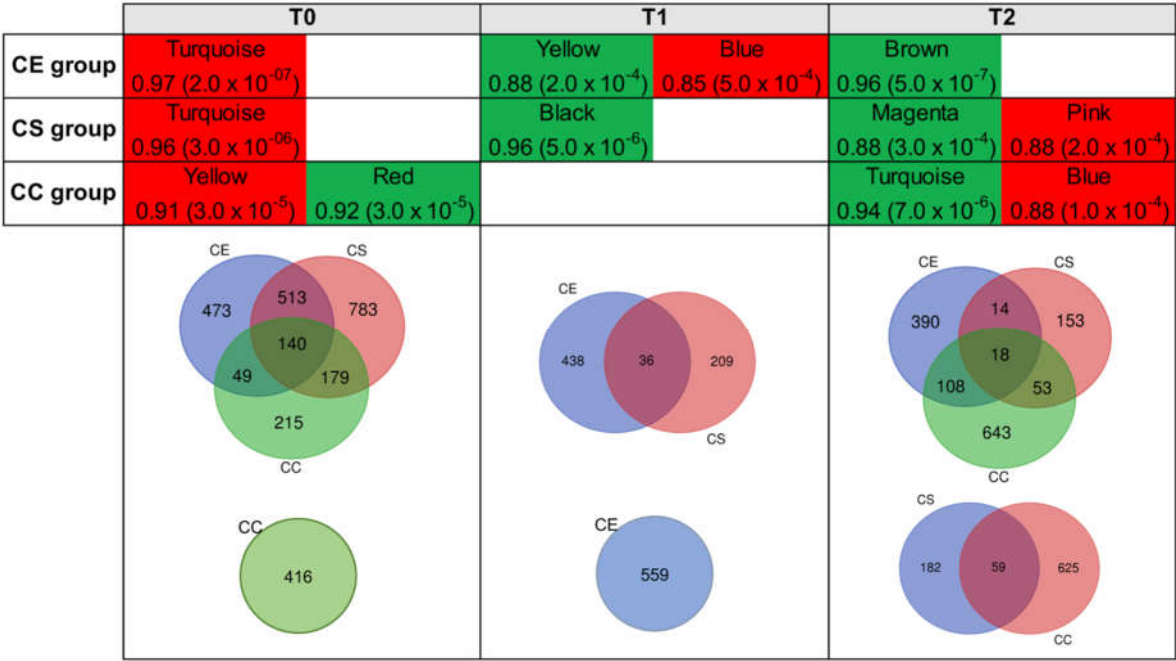


Figure 5. Highly correlated modules ($MM > 0.85$) with time-interval found for the CE, CS, and CC groups. The box shows the MS, module significance value, i.e. GS mean for module. The color bar indicates the MS value range, which red or green are positive or negative values, respectively. The boxes colored in red or green indicate, respectively, positive or negative correlation with a time interval. Venn diagram analysis for module gene composition by time-interval.

Subsequently, we performed KEGG enrichment on these module genes and terms related to metabolism, pathways in cancer, signaling pathways, protein transport, and transcription were found for the three groups (Spreadsheet S1). Additionally, terms related to cellular response to bacterial infection were found in the module positively correlated with the T0 interval for CE and CS groups. In T1, various terms associated with immune response and autophagy were over-represented on modules correlated to group CE. In T2 terms involved in signaling pathways related to immune response and ubiquitination were found for the group CE. In the group CS, only a few terms related to the cellular response to bacterial infection were found.

The comparative analysis of the top 20 highest-ranked KEGG terms found for CE, CS, and CC groups by time-interval showed that these three groups presented more similarity among them in the T0 interval than in the T1 or T2 intervals (Fig. 6, Spreadsheet S1). Interestingly, we found specific cellular processes for each group per time-interval. We describe below the most relevant terms, i.e., pathways related to cellular response to bacterial infection. In T0 and positively correlated modules, the group CE has terms as mitophagy and pathogenic *E. coli* infection. The intersection area between CE and CS groups presented terms as peroxisome and legionellosis. In T1 and negatively correlated modules, four terms such as Fc gamma R-mediated phagocytosis, Salmonella infection, Notch signaling pathway, and T-cell receptor signaling pathway appear in the group CE. For group CS, two over-represented terms were found: ErbB signaling pathway and pathogenic *E. coli* infection. Additionally, only in the group CE in this interval presented positively correlated module - blue - and enriched for terms such as autophagy and lysosome. In T2 interval and negatively correlated modules, the group CE presented many terms related to immune response such as hippo, FoxO, VEGF, and p53 signaling pathways, and related to bacterial infection as Shigellosis. In contrast, the group CS had two terms related to bacterial infection: ferroptosis and TGF-beta signaling pathway. In this interval CS and

CC groups presented modules, respectively pink and blue, positively correlated but none of the top 20 over-represented terms was in the intersection area (Spreadsheet S1).

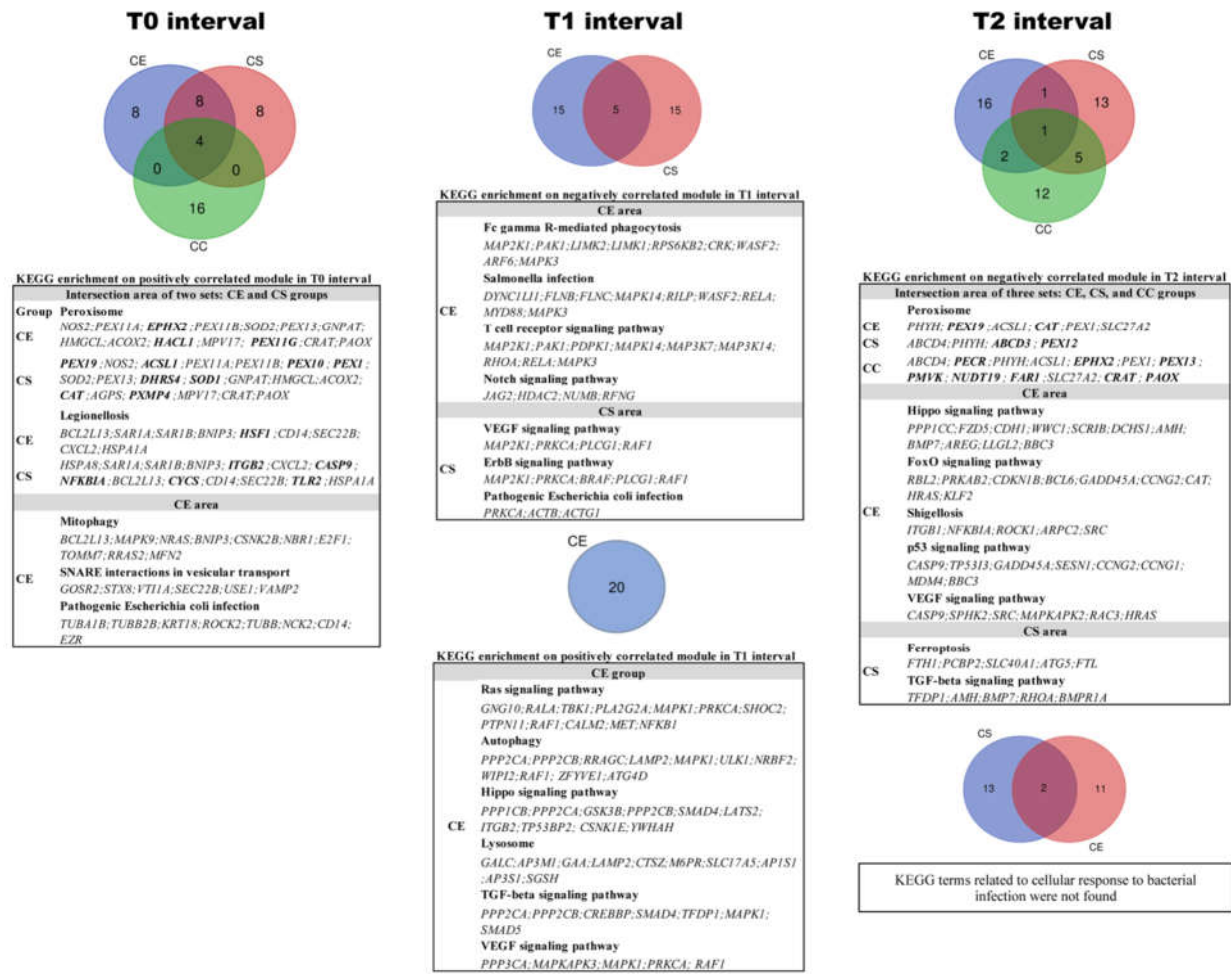


Figure 6. Comparative analysis of the top 20 KEGG terms over-represented on module genes highly correlated with each time-interval for the CE, CS, and CC groups. Venn diagram obtained for the T0, T1, and T2 time-intervals. The boxes show the KEGG terms involved in cellular response to bacterial infection found in each area of the diagram. Genes in bold letters indicate genes in the specific pathway for the group.

3.3. Co-expression network of the genes involved in autophagy, lysosome, or signaling pathways

We selected 43 genes involved in autophagy, lysosome, or signaling pathways belonging to the blue module – unique highly and positively correlated to T1 in CE group – of the CE network (Fig. 6). The gene expression data obtained for the T1 interval was used for network construction. For comparative propose we also constructed for CS group. The Figure 7 shows the CE and CS subnetworks with same gene-gene link threshold (Pearson's $r > 0.85$). The CE and CS subnetworks presented different topologies. In CE subnetwork all 43 genes were strongly interconnected. While in CS subnetwork only 20 genes were connected with one or two genes each and in the WGCNA network these genes belong to different modules. Through this result we can confirm that autophagy – including lysosome and signaling pathways involved in this process – is the distinctive cellular response between EIEC and *S. flexneri* after one hour of bacteria-cell infection.

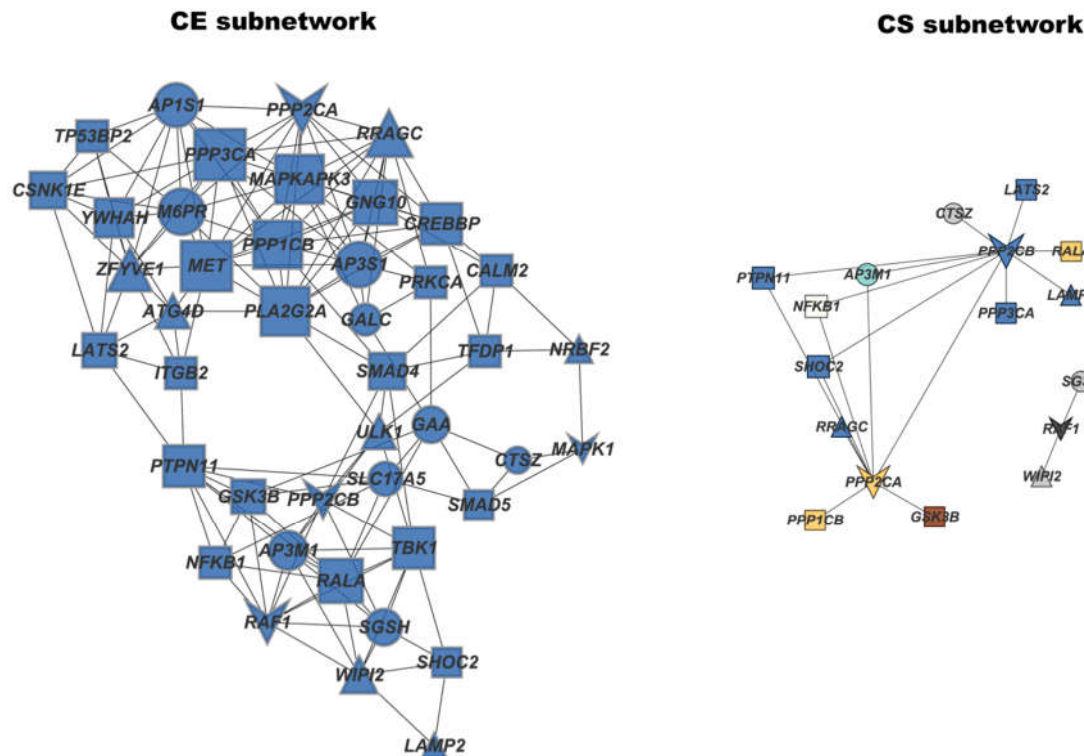


Figure 7. CE and CS subnetworks for autophagy, lysosome, and signaling pathways genes. Node fill colors indicate the module in the CE or CS WGCNA networks. Triangle, vee, rectangle, and circle nodes indicates respectively genes involved in autophagy, autophagy/signaling, signaling, or lysosome. Node size is proportional to the number of gene-gene links. Link strength threshold > 0.85.

4. Discussion

Our first question was how different are EIEC and *S. flexneri* (M90T) during bacterial infection of Caco-2 cells. Previous studies have shown that EIEC comparatively with *S. flexneri* has a lower expression of *icsA* and *icsB* when infecting epithelial or macrophagic cells [7, 11]. Here, it was found that these two genes were not differentially expressed between these two bacteria in Caco-2 conditioned medium. Since, the differential expression of virulence genes is not sufficient for explaining the lower virulence of EIEC, here we conducted a comparative transcriptomic and gene co-expression network analyses for further investigate the differential virulence of genetically close pathogenic *E. coli* [20]. The differential gene expression analysis revealed that 12% of the transcribed genes were DE genes (chromosomal, pINV, and foreign/horizontal acquired genes) between EIEC and *S. flexneri*, which the majority of them hypo-expressed in EIEC.

The gene co-expression network analysis showed that EIEC and *S. flexneri* networks presented different topologies. The EIEC network revealed that pINV genes are not integrated within the major network. Conversely, in the *S. flexneri* network many pINV genes are connected with other chromosomal genes and the majority of them are hubs. These hubs, belonging to the *ipa* family, are involved in bacterial invasion (*ipas*) and in cell-to-cell spread (*spa* and *icsB*) [4, 15-19].

Our results indicate that not only the hyper- or hypo-expression of the pINV genes but also the pINV genes functional cooperation with chromosomal genes confer the higher virulence of *S. flexneri*. This “cross-talking” regulation of plasmid and chromosome genes reduces fitness costs and could confer advantageous traits, such as antibiotic resistance or pathogenesis [32]. However, the “inappropriate” expression of acquired genes may also

lead to a loss of bacterial competitive fitness, as pointed by Vial and Hommais [32]. We found that EIEC presents a higher number of exclusively expressed acquired genes than *S. flexneri*. Notwithstanding, it was observed a diminished number of intracellular EIEC, as compared with *S. flexneri*, during cell-to-cell bacterial dissemination [11].

Our second question was how the enterocytes “see” the EIEC and the *S. flexneri* during the three-hour cellular infection period. WGCNA and eigengene analysis reflected the enterocyte global gene expression variation along this period. Additionally, in the T1 interval EIEC induced a higher number of gene expression changes in Caco-2 cells than *S. flexneri*. This variation seems to be associated with gene dysregulation induced by the bacteria or, alternatively, to adaptive cellular processes in response to bacterial infection. The CC group also presented gene expression variation, probably due to the *in vitro* experimental stress. Some studies have shown that topological alteration in gene co-expression networks constitute a pressure for adapting the cellular metabolism from one steady-state to another, frequently in response to environmental changes [21, 33]. It is known that *S. flexneri* promotes gene dysregulation and transcription remodeling, inducing metabolic changes in the host-cells for bacterial survival and replication [34, 35].

To investigate the molecular mechanism involved in the cellular response to bacterial invasion we identified transcriptional modules highly correlated with time-interval for CE, CS, or CC groups. A transcriptional module encompasses genes presenting high co-variation expression and its expression profile is distinct from the other modules in the co-expression network [28]. Hence, transcriptional modules correspond to a group of genes functionally linked, which play specific biological functions and/or belong to a common molecular pathway [21, 28]. The dynamic functional analysis showed that in T0, Caco-2 infected with EIEC or *S. flexneri* exhibited many genes involved in peroxisome and several genes were common to CS and CE groups (Fig. 5). However, after one-hour post-infection EIEC induced a higher number of genes involved in host defense than *S. flexneri*. Several genes involved in immune response, such as autophagy, lysosome, and signaling pathways only appear in the CE group (Fig. 5 and Table 5). Moreover, all these genes present high expression covariation, indicating that they are linked to specific biological functions or pathways, i.e., autophagy (Fig. 6). A previous study showed that EIEC intracellular multiplication and/or cell-to-cell spreading were diminished when compared with *S. flexneri* [11].

The transcriptional module analysis showed that autophagy and signaling pathways related to the immune response after one-hour EIEC post-infection are probably associated with bacterial clearance. A variety of intracellular bacterial pathogens, such as *Shigella*, *Salmonella*, *Listeria*, *Mycobacteria* and *Legionella* can interact with the autophagy machinery, and part of these pathogens have evolved mechanisms for evading from autophagy, minimizing its efficiency as an innate immune response to these infections [36-39].

Our results indicate that EIEC is inefficient for surviving inside Caco-2 cells when compared with *S. flexneri*. The “cross-talking” regulation of plasmid and chromosomal genes – as described above – is still evolving in EIEC. The bacterial co-expression network analysis showed that the pINV genes in EIEC are not connected with the chromosomal genes in the network, particularly *spa* and *icsB* genes, which are hub genes in the *S. flexneri* network. The *icsB* protein has been shown to be essential for *S. flexneri* survival, subverting autophagy recognition in the course of bacterial infection [17, 19, 40-43].

5. Conclusions

This study showed that the lower virulence of EIEC is probably associated to lack of functional cooperation between pINV and chromosomal genes, differently of what was observed in *S. flexneri*. Consequently, EIEC becomes less efficient in subverting host-cell bacterial recognition as well as defense mechanisms such as autophagy.

Supplementary Materials: **Figure S1.** WGCNA for CE group; **Figure S2.** WGCNA for CS group; **Figure S3.** WGCNA for CC group; **Table S1.** Biological functions of the differentially expressed

genes (DE) and genes expressed only in EIEC (E-EE) or *S. flexneri* (S-EE).; **Table S2.** Genes included for EIEC or *S. flexneri* co-expression network construction.; **Spreadsheet S1.** KEGG enrichment analysis on module genes highly correlated with each time-intervals for CE, CS, and CC groups.

Author Contributions: Conceptualization, Silvia Bando, Ana Moreno and Marina Martinez; Data curation, Silvia Bando and Fernanda Bertonha; Formal analysis, Silvia Bando, Ana Moreno, Hadassa Santos, Fernanda Bertonha and Lucas Ferreira; Funding acquisition, Marina Martinez and Carlos Moreira-Filho; Investigation, Silvia Bando, Ana Moreno, Hadassa Santos and Fernanda Bertonha; Methodology, Silvia Bando, Ana Moreno, Hadassa Santos, Fernanda Bertonha and Lucas Ferreira; Resources, Marina Martinez and Carlos Moreira-Filho; Visualization, Silvia Bando and Fernanda Bertonha; Writing – original draft, Silvia Bando and Ana Moreno; Writing – review & editing, Silvia Bando, Ana Moreno, Fernanda Bertonha, Marina Martinez and Carlos Moreira-Filho.

Funding: This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) grants 2015/22308-2 to CAM-F and 2015/16505-0 to ACRM; by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant no. 306893/2018-5 to CAM-F; ACRM received a fellowship (2016/0078-1) from FAPESP.

Data Availability Statement: All microarray raw data were deposited in GEO public database (www.ncbi.nlm.nih.gov/geo), under accession numbers: GSE93588, GSE93525, and GSE93587 for Caco-2 cells; raw data for bacterial gene expression arrays were deposit (Feb 23th, 2022) in GEO and accession numbers will be provided.

Acknowledgments: The *S. flexneri* strain M90T was gently provided by Dr. Phillip Sansonetti, Institut Pasteur, France. We would like to thank the skillful technical assistance of Leandro Rodrigues Ferreira.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Organization WH. Diarrhoeal Disease 2017 [cited 2022 Feb/21]. Available from: <https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease>.
2. Bellido-Blasco J, Arnedo-Pena A. Epidemiology of Infectious Diarrhea. In: Nriagu J, editor. Encyclopedia of Environmental Health. Second ed: Elsevier; 2011. p. 659-71.
3. Moreno AC, Filho AF, Gomes ToA, Ramos ST, Montemor LP, Tavares VC, et al. Etiology of childhood diarrhea in the northeast of Brazil: significant emergent diarrheal pathogens. *Diagn Microbiol Infect Dis*. 2010;66(1):50-7. doi: 10.1016/j.diagmicrobio.2008.03.017.
4. Parsot C. Shigella spp. and enteroinvasive Escherichia coli pathogenicity factors. *FEMS Microbiol Lett*. 2005;252(1):11-8. doi: 10.1016/j.femsle.2005.08.046.
5. Escobar-Páramo P, Giudicelli C, Parsot C, Denamur E. The evolutionary history of Shigella and enteroinvasive Escherichia coli revised. *J Mol Evol*. 2003;57(2):140-8. doi: 10.1007/s00239-003-2460-3.
6. DuPont HL, Formal SB, Hornick RB, Snyder MJ, Libonati JP, Sheahan DG, et al. Pathogenesis of Escherichia coli diarrhea. *N Engl J Med*. 1971;285(1):1-9. doi: 10.1056/NEJM197107012850101.
7. Bando SY, Moreno AC, Albuquerque JA, Amhaz JM, Moreira-Filho CA, Martinez MB. Expression of bacterial virulence factors and cytokines during in vitro macrophage infection by enteroinvasive Escherichia coli and Shigella flexneri: a comparative study. *Mem Inst Oswaldo Cruz*. 2010;105(6):786-91. doi: 10.1590/s0074-02762010000600009.
8. Lan R, Alles MC, Donohoe K, Martinez MB, Reeves PR. Molecular evolutionary relationships of enteroinvasive Escherichia coli and Shigella spp. *Infect Immun*. 2004;72(9):5080-8. doi: 10.1128/IAI.72.9.5080-5088.2004.
9. Adriana Gibotti A, Tanaka T, Oliveira V, Taddei C, Martinez M. Molecular characterization of enteroinvasive *Escherichia coli ipa* genes by PCR-RFLP analysis. *Brazilian Journal of Microbiology*. 2004;35:1-2. doi: 10.1590/S1517-83822004000100012.

10. Moreno AC, Ferreira KS, Ferreira LG, Almeida SR, Martinez MB. Recognition of enteroinvasive *Escherichia coli* and *Shigella flexneri* by dendritic cells: distinct dendritic cell activation states. *Mem Inst Oswaldo Cruz*. 2012;107(1):138-41. doi: 10.1590/s0074-02762012000100020.
11. Moreno AC, Ferreira LG, Martinez MB. Enteroinvasive *Escherichia coli* vs. *Shigella flexneri*: how different patterns of gene expression affect virulence. *FEMS Microbiol Lett*. 2009;301(2):156-63. doi: 10.1111/j.1574-6968.2009.01815.x.
12. Harris JR, Wachsmuth IK, Davis BR, Cohen ML. High-molecular-weight plasmid correlates with *Escherichia coli* enteroinvasiveness. *Infect Immun*. 1982;37(3):1295-8. doi: 10.1128/iai.37.3.1295-1298.1982.
13. Sansonetti PJ, Kopecko DJ, Formal SB. Involvement of a plasmid in the invasive ability of *Shigella flexneri*. *Infect Immun*. 1982;35(3):852-60. doi: 10.1128/iai.35.3.852-860.1982.
14. Silva RM, Toledo MR, Trabulsi LR. Correlation of invasiveness with plasmid in enteroinvasive strains of *Escherichia coli*. *J Infect Dis*. 1982;146(5):706. doi: 10.1093/infdis/146.5.706.
15. Jain S, van Ulsen P, Benz I, Schmidt MA, Fernandez R, Tommassen J, et al. Polar localization of the autotransporter family of large bacterial virulence proteins. *J Bacteriol*. 2006;188(13):4841-50. doi: 10.1128/JB.00326-06.
16. Bourdet-Sicard R, Egile C, Sansonetti PJ, Tran Van Nhieu G. Diversion of cytoskeletal processes by *Shigella* during invasion of epithelial cells. *Microbes Infect*. 2000;2(7):813-9. doi: 10.1016/s1286-4579(00)90366-6.
17. Egile C, Loisel TP, Laurent V, Li R, Pantaloni D, Sansonetti PJ, et al. Activation of the CDC42 effector N-WASP by the *Shigella flexneri* IcsA protein promotes actin nucleation by Arp2/3 complex and bacterial actin-based motility. *J Cell Biol*. 1999;146(6):1319-32. doi: 10.1083/jcb.146.6.1319.
18. Kayath CA, Hussey S, El hajjami N, Nagra K, Philpott D, Allaoui A. Escape of intracellular *Shigella* from autophagy requires binding to cholesterol through the type III effector, IcsB. *Microbes Infect*. 2010;12(12-13):956-66. doi: 10.1016/j.micinf.2010.06.006.
19. Ogawa M, Yoshimori T, Suzuki T, Sagara H, Mizushima N, Sasakawa C. Escape of intracellular *Shigella* from autophagy. *Science*. 2005;307(5710):727-31. doi: 10.1126/science.1106036.
20. Bando SY, Iamashita P, Guth BE, Dos Santos LF, Fujita A, Abe CM, et al. A hemolytic-uremic syndrome-associated strain O113:H21 Shiga toxin-producing *Escherichia coli* specifically expresses a transcriptional module containing *dicA* and is related to gene network dysregulation in Caco-2 cells. *PLoS One*. 2017;12(12):e0189613. doi: 10.1371/journal.pone.0189613.
21. Bando SY, Iamashita P, Silva FN, Costa LDF, Abe CM, Bertonha FB, et al. Dynamic Gene Network Analysis of Caco-2 Cell Response to Shiga Toxin-Producing. *Microorganisms*. 2019;7(7). doi: 10.3390/microorganisms7070195.
22. Lea T. Caco-2 Cell Line. In: Verhoeckx KC, P López-Expósito, I Kleiveland, C Lea, T Mackie, A Requena, T Swiatecka, D Wichers, H, editor. *The impact of food bioactives on health: in vitro and ex vivo models*. Cham: Springer; 2015. p. 103-11.
23. Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.
24. Smyth G. *Limma: linear models for microarray data*. Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W, editors. New York: Springer; 2005.
25. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-504. doi: 10.1101/gr.1239303.
26. 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. doi: 10.1371/journal.pone.0128174.

27. 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. doi: 10.1371/journal.pone.0079913.
28. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559. doi: 10.1186/1471-2105-9-559.
29. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013;14:128. doi: 10.1186/1471-2105-14-128.
30. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;44(W1):W90-7. doi: 10.1093/nar/gkw377.
31. Calculate and Draw Custom Venn Diagrams. *Bioinformatics and Evolutionary Genomics*. GHENT University, Belgium. Available from: <http://bioinformatics.psb.ugent.be/webtools/Venn/>.
32. Vial L, Hommais F. Plasmid-chromosome cross-talks. *Environ Microbiol*. 2020;22(2):540-56. doi: 10.1111/1462-2920.14880.
33. Sieberts SK, Schadt EE. Moving toward a system genetics view of disease. *Mamm Genome*. 2007;18(6-7):389-401. doi: 10.1007/s00335-007-9040-6.
34. Killackey SA, Sorbara MT, Girardin SE. Cellular Aspects of Shigella Pathogenesis: Focus on the Manipulation of Host Cell Processes. *Front Cell Infect Microbiol*. 2016;6:38. doi: 10.3389/fcimb.2016.00038.
35. Sarantis H, Grinstein S. Subversion of phagocytosis for pathogen survival. *Cell Host Microbe*. 2012;12(4):419-31. doi: 10.1016/j.chom.2012.09.001.
36. Sil P, Muse G, Martinez J. A ravenous defense: canonical and non-canonical autophagy in immunity. *Curr Opin Immunol*. 2018;50:21-31. doi: 10.1016/j.coi.2017.10.004.
37. Jo EK, Yuk JM, Shin DM, Sasakawa C. Roles of autophagy in elimination of intracellular bacterial pathogens. *Front Immunol*. 2013;4:97. doi: 10.3389/fimmu.2013.00097.
38. Cadwell K. Crosstalk between autophagy and inflammatory signalling pathways: balancing defence and homeostasis. *Nat Rev Immunol*. 2016;16(11):661-75. doi: 10.1038/nri.2016.100.
39. Kimmey JM, Stallings CL. Bacterial Pathogens versus Autophagy: Implications for Therapeutic Interventions. *Trends Mol Med*. 2016;22(12):1060-76. doi: 10.1016/j.molmed.2016.10.008.
40. Weddle E, Agaisse H. Spatial, Temporal, and Functional Assessment of LC3-Dependent Autophagy in Shigella flexneri Dissemination. *Infect Immun*. 2018;86(8). doi: 10.1128/IAI.00134-18.
41. Leung Y, Ally S, Goldberg MB. Bacterial actin assembly requires toca-1 to relieve N-wasp autoinhibition. *Cell Host Microbe*. 2008;3(1):39-47. doi: 10.1016/j.chom.2007.10.011.
42. Baxt LA, Goldberg MB. Host and bacterial proteins that repress recruitment of LC3 to Shigella early during infection. *PLoS One*. 2014;9(4):e94653. doi: 10.1371/journal.pone.0094653.
43. Campbell-Valois FX, Sachse M, Sansonetti PJ, Parsot C. Escape of Actively Secreting Shigella flexneri from ATG8/LC3-Positive Vacuoles Formed during Cell-To-Cell Spread Is Facilitated by IcsB and VirA. *MBio*. 2015;6(3):e02567-14. doi: 10.1128/mBio.02567-14.