

Screening for host proteins interacting with *Escherichia coli* O157:H7 EspF using bimolecular fluorescence complementation

Ying Hua^{‡,1,2}, Jingwei Ju^{‡,3,4}, Xiangyu Wang^{1,2}, Bao Zhang^{1,2}, Wei Zhao^{1,2}, Qiwei Zhang^{1,2}, Yingzhu Feng⁵, Wenbin Ma^{*,3,4} & Chengsong Wan^{**1,2}

¹Biosafety Level 3 Laboratory, School of Public Health, Southern Medical University, Guangzhou 510515, China

²Key Laboratory of Tropical Disease Research of Guangdong Provincial, Guangzhou 510515, China

³Key Laboratory of Gene Engineering of the Ministry of Education, State Key Laboratory for Biocontrol, School of Life Sciences, Sun Yat-sen University, Guangzhou 510006, China

⁴Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University, Guangzhou 510006, China

⁵Guangzhou Institutes of Biomedicine & Health, Chinese Academy of Sciences, Guangzhou 510530, China

* Author for correspondence: mawenbin@mail.sysu.edu.cn

** Author for correspondence: gzwcs@smu.edu.cn

‡ Authors contributed equally

Aim: To screen host proteins that interact with enterohemorrhagic *Escherichia coli* O157:H7 EspF. **Materials & methods:** Flow cytometry and high-throughput sequencing were used to screen interacting proteins. Molecular function, biological processes and Kyoto Encyclopedia of Genes and Genomes pathways were studied using the DAVID online tool. Glutathione S-transferase pull down and dot blotting were used to verify the interactions. **Results:** 293 host proteins were identified to associate with EspF. They were mainly enriched in RNA splicing ($p = 0.005$), ribosome structure ($p = 0.012$), and involved in 109 types of signaling pathways. SNX9 and ANXA6 were confirmed to interact with EspF. **Conclusion:** EspF interacts with ANXA6; they may form a complex to manipulate the process of phagocytosis; EspF plays a highlighted pathogenic role in enterohemorrhagic *E. coli* infection process.

First draft submitted: 15 May 2017; Accepted for publication: 1 September 2017; Published online: 11 December 2017

Keywords: ANXA6 • BiFC • bioinformatics analysis • coprecipitation • enterohemorrhagic *Escherichia coli* O157:H7 • EspF • flow sorting • KRAS • protein–protein interactions • subcellular localization

Enterohemorrhagic *Escherichia coli* (EHEC) O157:H7 is an important food-borne causative agent in sporadic outbreaks of illness such as diarrhea, hemorrhagic colitis and hemolytic–uremic syndrome. Severe symptoms may even lead to death [1–4]. Research into EHEC has mainly explored the development of attaching and effacing (A/E) intestinal lesions and the secretion of Shiga toxin in EHEC infection [5–7]. The protein EspF is one of the multifunctional effectors of A/E lesions induced by EHEC and enteropathogenic *E. coli* (EPEC). EspF participates in a number of damage processes in host cells, targeting mitochondria and nucleoli [6,8] and disrupting the tight junctions (TJs) of intestinal epithelial cells [9], leading to cytoskeletal rearrangements, actin polymerization and pedestal formation [10]. EspF thus emerges as the ‘Swiss army knife’ of a bacterial pathogen [11,12]. EHEC O157:H7 employs a type-III secretion system to export toxic factors and effector proteins to host cells after adhering to the brush border of epithelial cells [5,13–14]. The production of virulence factors, subsequent invasion and diffusion, and the interaction of effector proteins with host proteins are key steps in EHEC O157:H7 pathogenesis. The mechanism by which EspF breaks through the intestinal epithelial barrier into host cells, the host proteins that EspF interacts with, and how cellular damage, and even apoptosis, result remain unclear. Studying the pathogen–host interaction process will increase the emerging knowledge about EHEC O157:H7 pathogenesis.

The EspF protein has been shown to induce apoptosis after ‘injection’ into host cells [1]. The N-terminal (1–73 amino acids [aa]) of this approximately 27 kDa protein contains a secretory signal (1–20 aa), a mitochondrial-targeting signal (1–24 aa) and a nucleolar-targeting domain (21–74 aa); 73–248 aa consists of four proline-rich repeats (PRRs), each containing a eukaryotic cell sorting nexin 9 (SNX9) protein-binding site SH3 motif, an

efficient CRIB domain, and a possible actin-binding domain [11]. Such structural features enable EspF to target the mitochondria, resulting in disruption of the mitochondrial membrane potential, release of cytochrome c into the cytoplasm, cleavage of caspases 9 and 3; all of these events are associated with mitochondrial apoptotic signaling pathways [15,16]. However, the SH3 and CRIB domains enable EspF to interact with the SNX9 protein and the neuronal Wiskott–Aldrich syndrome (N-WASP) protein directly [17,18]. Moreover, VK Viswanathan *et al.* have shown that EspF can interact with the host intermediate filament protein cytokeratin 18 (CK18) in a complex with adaptor protein 14-3-3, which can alter the architecture of the intermediate filament network in EPEC-infected cells [19]. It cannot only modulate the structure and function of TJs but can also remodel the brush border [20,21]. In addition, EspF can inhibit EPEC internalization by J774.A1 macrophages through a PI-3 kinase-dependent pathway [22].

EspF is well recognized as a significant effector protein, but little research has been done on its interaction with host proteins. The identification of bacteria–host interactions is therefore crucial in elucidating the pathogenic mechanism. Bimolecular fluorescence complementation (BiFC) analysis is a new technique to determine the localization and interaction of target proteins in living cells, first reported by Hu *et al.* in 2002 [23]. In the BiFC assay, protein–protein interactions bring two fragments of a fluorescent protein (tagged to two separate proteins) within close proximity and allow for their cofolding into a functional fluorescent protein [24]. Using BiFC analysis technology and the hORFeome human cDNA library [25], we tethered YFPn (yellow fluorescent protein, 1–155 aa) to the C-terminal ends of the EspF protein to form EspF-CVN, constructed a stable monoclonal cell line expressing YFPc (156–239 aa) and then transferred EspF-CVN to the stable monoclonal cell line to construct a bistable cell line. Screening of the fluorescence signals of positive cells was carried out using flow cytometry. After high-throughput sequencing, the exogenous cDNA gene, which has a positive interaction with EspF, was identified by primer tagging resulting in a list of proteins that directly interact with EspF. Protein annotation analysis was performed using a bioinformatics method (The Database for Annotation, Visualization and Integrated Discovery [DAVID], Gene Ontology [GO] and Kyoto Encyclopedia of Genes and Genomes [KEGG]). The GO functional annotation used included the molecular function (MF), biological process (BP) and cellular component (CC) of the related proteins. The STRING online tool was used to analyze interactions between target proteins, and Cytoscape software was used to draw the protein interaction network. Functional annotation and pathway analysis allowed the exploration of functional enrichment trends, the signaling pathways involved and the interaction network, providing new ideas and methods for studying the interaction between pathogen and host proteins. Subcellular localization of EspF was performed by confocal microscopy, and we conducted a verification of the screened proteins by coprecipitation. Our research provides a solution map of proteins that interact with EspF and lays a foundation for further study of the EspF–host cell apoptosis pathogenic process.

In this study, we identified a series of host proteins interacting with EspF, using a genomics approach. Bioinformatics analysis of the obtained proteomics data revealed that EspF protein may employ a huge protein–protein interaction network to induce colitis and even colorectal carcinogenesis.

Materials & methods

Cell lines & bacterial strains

293T is a human renal epithelial (293) cell line transfected with the adenovirus E1A gene; plasmids containing the SV40 origin of replication and the promoter region can be replicated in this cell line, which is mainly used for exogenous protein overexpression. BOSC23 cells are modified from 293T and mainly used for packaging viruses. HTC75 cells are human fibrosarcoma cells. Bacterial strains EHEC O157:H7 EDL 933 were obtained (Center for Disease Control and Prevention, China) and *E. coli* DH5 α was preserved in our laboratory.

Establishment of retroviral pool screening libraries

We performed Individual Gateway recombination reactions (in 96-well plates) for all 18,000 open-reading frames (ORFs) from the hORFeome v7.1 library [26] with a mixture of pBabe-CMV-YFPc-DEST-puro (pB-CMV-NC-puro) and pBabe-CMV-DEST-YFPc-puro (pB-CMV-CC-puro; 1:1; Invitrogen, CA, USA) to generate ORFs tagged with YFPc at either the N or C terminus, obtaining a library of YFPc-prey pool plasmids.

The YFPc-prey collections were then transfected into cells to generate retroviruses. BOSC23 cells were plated on a 10-cm dish. YFPc-prey, PCGP and VSVG packaging plasmids (3:2:1; Invitrogen) were diluted with Opti minimal essential medium (Opti-MEM; Gibco, CA, USA) and mixed with polyethylenimine (PEI; Polysciences, FL, USA). Added to BOSC cells after 20 min, they were then incubated in 5% CO₂ at 37°C for routine passage.

The culture medium was replaced by fresh medium 6 h after transfection, and the virus supernatant was collected at 48 h and stored at -80°C . This process resulted in a database of 36 viruses.

Establishment of EspF–YFPn stable cell line

The EHEC O157: H7 EspF ORF was amplified by the PCR and cloned into the vector pENTR-D-TOPO (Invitrogen). The target gene *espF* was subcloned into a retroviral expression vector, pBabe-CMV-DEST-YFPn-neo(pB-CMV-CVN-neo), fused with YFPn at its C terminal. The eukaryotic expression vectors of the EspF and YFPn fusion protein were obtained by transforming the DH5 α competent cells and extracting the monoclonal colonies and plasmids (Macherey-Nagel, Düren, Germany).

As described above, the virus was packaged in BOSC cells with EspF–YFPn plasmid, PSPAX2 and PMD2G packaging plasmid (4:2:1; Invitrogen). After transfection for 6 h, the medium was replaced with fresh medium and the virus supernatant was collected 48 h later.

The day before the virus infection, HTC75 cells were plated in 6-well plates. The virus solution was carefully extracted with a 5-ml syringe, filtered through a 0.45- μm filter and added to HTC75 cells. Polybrene (Sigma-Aldrich, Shanghai, China) was then added to a final concentration of 4 $\mu\text{g}/\text{ml}$. After 12–24 h of virus infection, the medium was replaced by Dulbecco's modified Eagle medium with 10% (vol/vol) fetal bovine serum (FBS; Gibco). 48 h after infection, G418 (Sigma-Aldrich) was added to a final concentration of 500 $\mu\text{g}/\text{ml}$ for screening. The medium was replaced with fresh culture medium when the control group cell death was 100% after 7 days of screening, until a stable HTC75 cell line expressing the EspF–YFPn plasmid was obtained.

Construction of a BiFC (YFPc & EspF–YFPn library) coexpression stable cell line

Bait protein stably expressing cells with YFPn were infected with the 36 virus libraries obtained as described above. 48 h after infection, puromycin (1 $\mu\text{g}/\text{ml}$; Sigma-Aldrich) was added to obtain YFPn–EspF and YFPc–X double stable transfection cell lines.

Sorting of positive cells

The 'YFPc and EspF–YFPn library' double stably transfected HTC75 cell cultures were expanded to the appropriate scale, then digested with 0.05% Trypsin–ethylenediaminetetraacetic acid (Gibco) and centrifuged. The culture medium was discarded, and the cultures were resuspended with an appropriate volume of $1 \times$ PBS, added to the flow tube by a filter cap, and placed on a flow cytometry/cell sorter (BD FACSAria, BD Biosciences, NJ, USA). A flow tube containing fresh 10% FBS medium (HyClone, UT, USA) was also prepared to receive the sorted cells. The empty vector pBabe–CMV–YFPc–DEST–puro and pBabe–CMV–DEST–YFPc–puro were cotransfected into the HTC75 cell line and used as a negative control, which cannot interact; therefore, it cannot produce complementary fluorescence detectable by flow cytometry. Protein expression vectors pBabe–CMV–YFPc–POT1–puro and pBabe–CMV–TPP1–YFPc–puro were used as positive controls because these two proteins (POT1 and TPP1) have been shown to interact [27]. We carried out three rounds of sorting, and the positive control was used to adjust the parameters of the flow cytometer for the first sorting, and negative controls were included for each round. Only positive cells were received into each group. The cells were expanded and cultured for the next round of sorting. After several rounds of sorting, we obtained cells with a positive rate $>90\%$.

High-throughput sequencing

The positive cells were collected, and then their RNA was extracted (Invitrogen) and reverse transcribed to cDNA (BIO-RAD, CA, USA). PCR was carried out using cDNA as a template and specific primers that could recognize N-terminal and C-terminal sequences of pB–CMV–NC/CC–puro, with the PCR products subjected to DNA purity and complete assay. After detection, the PCR products were disrupted by an ultrasonic crusher (Covaris, MA, USA), and we constructed a library by end repairing, adding A-tail, adding sequencing adapter and conducting purification and PCR amplification (TaKaRa, Osaka, Japan). Following construction of the library, we applied the Q-PCR method to quantify the effective concentration of the library to ensure its quality. After library detection, we performed sequencing using the Illumina high-throughput sequencing platform (HiSeq/MiSeq; Beijing Novogene Technology, Beijing, China) [28,29], then utilized proteomic bioinformatics analysis (Bowtie2) to obtain the list of interacting proteins.

Bioinformatics analysis

The MFs, BPs, CCs, cell localization and KEGG signal pathways of proteins were analyzed using the DAVID online tool [30]. The DAVID database Homo Sapiens was used as a background reference, and we analyzed 293 proteins using GO and KEGG enrichment analysis. Fisher's exact test was used for statistical analysis and R-3.3.1 for statistical drawing. Protein interaction was analyzed with STRING online [31] and the protein interaction network mapped using Cytoscape_v3.4.0.

Subcellular localization

To determine the subcellular localization of the EspF protein *in vitro*, we used PCR to amplify *espF* and its N terminal (1–219) and C terminal (220–747) from *E. coli* EDL 933 genomic DNA using EasyPfu DNA polymerase (TransGen Biotech, Beijing, China). Primer pairs E-F: CCCAAGCTTGCCACCATGCTTAATGGAATTAGTAACGCT and E-R: CGCGGATCCCTTCTTCGATTGCTC were used to amplify the *espF* gene; primer pairs E-F and GFP-N-R: CGCGGATCCGAGTAAATGAAGTCAC were used to amplify *espF* 1–219; and primer pairs E-R and GFP-C-F: CCCAAGCTTGCCACCATGTCTCGTCCGGCACC GCCG were used to amplify *espF* 220–747 (the underlined portion indicates the target sequences of the restriction endonucleases *Bam*HI and *Hind*III). The PCR products were cloned into the *Bam*HI/*Hind*III restriction sites of pEGFP-N1 (Clontech, Osaka, Japan) to generate the mammalian expression vectors pEGFP–EspF, pEGFP–EspF/N and pEGFP–EspF/C. All constructs were verified by DNA sequencing.

HeLa cells were cultured following standard protocols. Cells were plated on confocal dishes (35 mm, NEST, Hong Kong, China) and allowed to grow overnight. Cells were transiently transfected with the mammalian expression vectors pEGFP, pEGFP–EspF, pEGFP–EspF/N and pEGFP–EspF/C by using the liposome transfection reagent Lipofectamine 3000 (Thermo Fisher Scientific, MA, USA) according to the manufacturer's instructions. 48 h after transfection, the transfected cells were stained with 250 nM MitoTracker Red (Beyotime, Shanghai, China) for 20 min in the dark and washed three-times with 1 × Hanks' Balanced Salt Solution solution (with Ca²⁺ & Mg²⁺). Then, the cells were fixed in 4% paraformaldehyde for 2 min at room temperature, the fixative was removed, and the nuclei in the monolayer were stained with DAPI (Beyotime) at room temperature for 5 min and washed three-times with 1 × Hanks' Balanced Salt Solution solution. The cells were examined under an FV1000 confocal microscope (Olympus, Tokyo, Japan).

Glutathione S-transferase pull down & dot blotting

To further validate the interaction between the bait protein EspF and the screened prey proteins, we chose some proteins randomly, and conducted coprecipitation by a glutathione S-transferase (GST) pull down. We constructed the prokaryotic expression vector pDEST–EspF (pDEST27; Invitrogen) and pMSCV–HA–Flag–prey (pMSCV–HA–Flag; Biovector) to express the bait protein EspF and candidate prey proteins tagged with GST and FLAG, respectively. We then cotransfected each bait-prey pair into 293T cells. After 48 h, we collected the cells, lysed them with 1 × NETN buffer (40 mM Tris-HCl, pH 8.0, 100 mM NaCl, 0.5% Nonidet P-40 and 1 mM EDTA, 10% glycerol), rotated them at 4°C for 15 min, centrifuged them at the highest speed for 10 min and transferred the supernatant to a new tube. Then, we added the supernatant to Glutathione Sepharose 4B beads (GE Healthcare Bio-Sciences AB, IL, USA) and incubated at 4°C for 2 h with gentle shaking. They were then centrifuged at 3000 r.p.m. for 5 min (4°C), the supernatant discarded and the cells washed with precooled 1 × NETN buffer for three-times. The bound proteins were then eluted into another new tube with elution buffer (50 mM Tris-HCl, pH 8.0, 20% glycerol, 20 mM reduced glutathione). Three microliters of each sample input and eluent were mixed with an equal volume of glycerol, plated on an appropriately sized nitrocellulose membrane on a 96-well plate (Corning, NY, USA), and then dropped on the hole and dried naturally. The samples were then incubated with blocking buffer (5% bovine serum albumin [BSA]), and the membrane was then incubated with anti-GST-horseradish peroxidase (HRP; cell signaling technology, MA, USA) and anti-FLAG–HRP (Sigma-Aldrich) antibodies.

Results

The average positive rate of the cell population interacting with EHEC O157:H7 EspF protein reached 98%

Using a flow-sorting approach, cells coexpressing YFP_n–EspF protein and YFP_c–X were detected by flow cytometry (BD FACSAria). Pairs of proteins known to interact with each other and pairs of proteins known to have no interaction served as positive and negative controls, respectively. BiFC-positive cells were then collected and

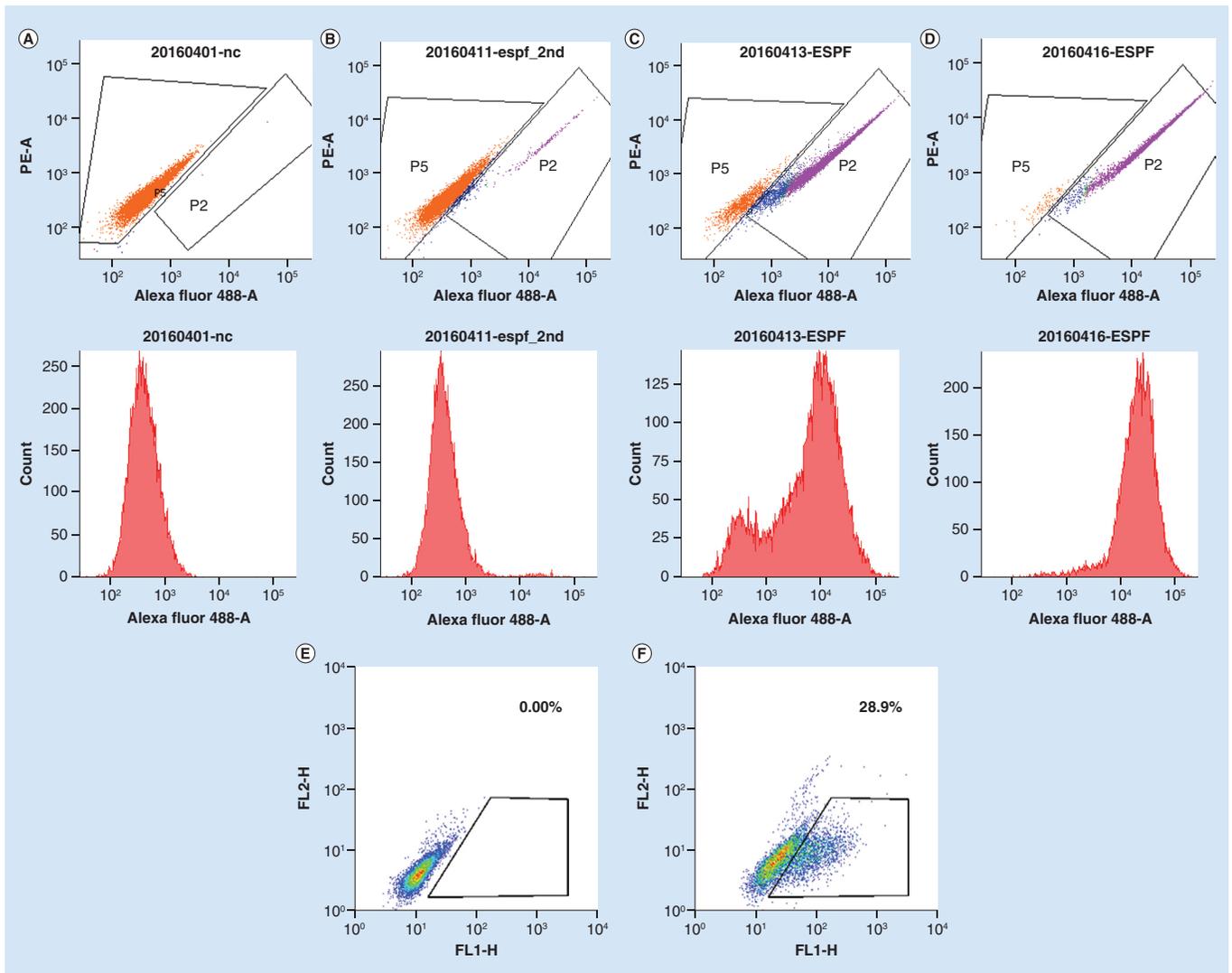


Figure 1. FACS analysis of fluorescence complementation in HTC75 cells expressing YFPn–EspF and YFPc-X. (A) Single-turn cDNA YFPc cell line used as negative control. **(B)** First round of EspF protein flow sorting. **(C)** Second round of EspF protein flow sorting. **(D)** Final result of EspF protein flow sorting. **(E)** HTC75 cells were detected as an empty control. **(F)** pBabe–CMV–YFPc–POT1–puro and pBabe–CMV–TPP1–YFPc–puro were cotransfected with HTC75 cells as a positive control, and the positive rate is 28.9%. Flow diagram shows that cells in the p2 box were collected each time. As rounds of selection were carried out, cell clusters migrated to the right and the positive bimolecular fluorescence complementation signal was enhanced, indicating that the interacting proteins were continuously enriched.

cultured. In the second round, we collected the positive cells for the next sorting. To reduce the signal-to-noise ratio and enhance the reliability of the results, the first and second rounds of sorting were repeated three-times, and only positive cells were collected each time. After three rounds of sorting, we obtained fluorescent cells containing reliable pairs of interacting proteins under natural conditions. At a positive rate $\geq 90\%$, the positive cells were considered true positives and the sorting process was discontinued.

As multiple rounds of screening progressed, the positive BiFC signal shifted to the right and increased, indicating the presence of interacting proteins that were continuously enriched (Figure 1). Figure 1A shows a single-turn cDNA YFPc cell line, as a flow-sorting negative control; Figure 1F shows the cotransfection of pBabe–CMV–YFPc–POT1–puro and pBabe–CMV–TPP1–YFPc–puro in HTC75 cells, as a positive control, and the positive rate is 28.9%; Figure 1D shows the final result of EspF protein flow sorting. The enrichment of sample cells was 97% accurate, and the enrichment of negative control cells was 1%, a difference from the positive detection rate of sample cells considered negligible.

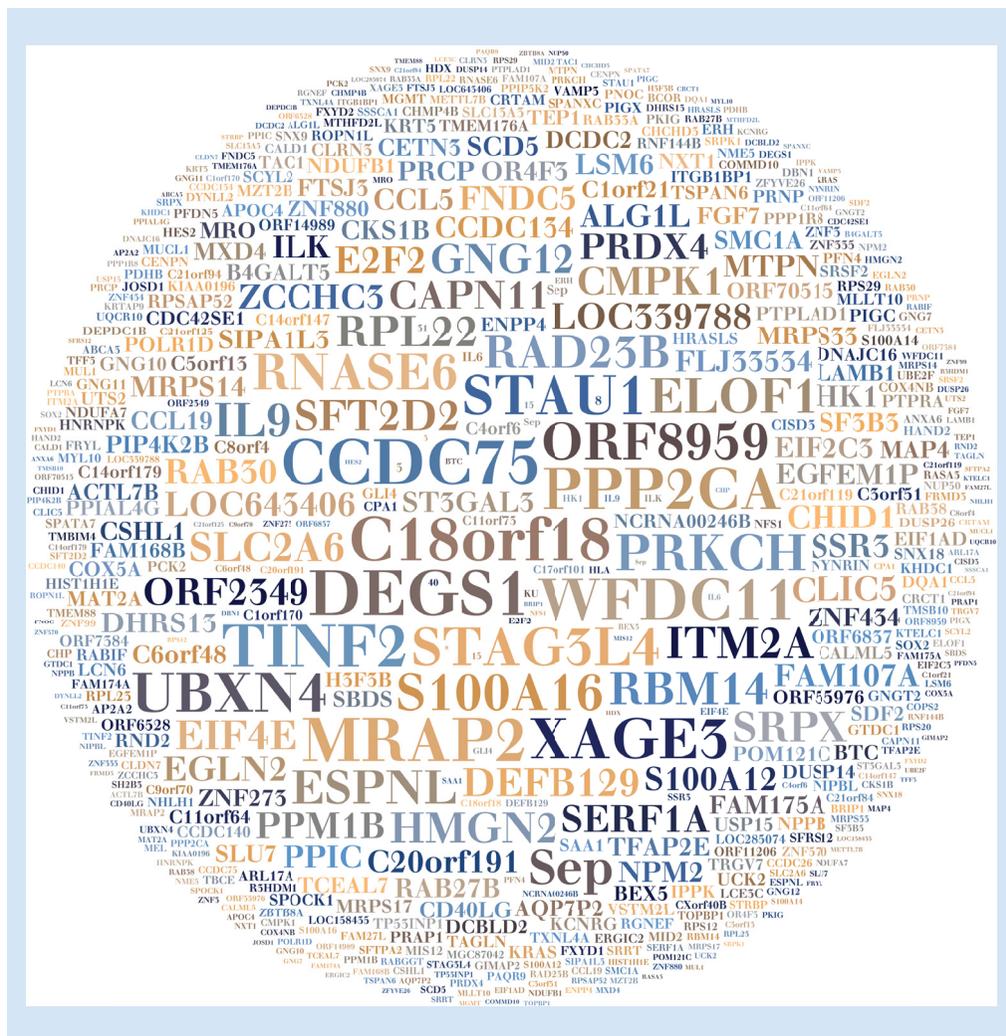


Figure 2. Cloud showing proteins that interact with enterohemorrhagic *Escherichia coli* O157: H7 EspF. A total of 293 proteins are displayed according to the reads configuration of protein sequencing. The greater the number of reads, the more prominently the protein is displayed on the cloud.

A total of 293 host cell proteins interacting with the EHEC O157:H7 EspF protein

We constructed a sequencing quality library, adding the sequencing primer binding sites to the ends of DNA. Following the first round of sequencing, the template strand of the first round was removed. To achieve sufficient quantity of template to conduct a second round of sequencing and complementary strand synthesis sequencing, we applied the paired-end module to guide the complementary strand in the original location to regeneration and amplification, effectively ensuring the accuracy of sequencing [32–34].

The readings were confirmed as meeting the tag and specific primer conditions. Bowtie 2 was used to align the short sequences to the gene sequence. The resulting SAM file was processed and converted with SAMtools; HTseq was then used to read the calculations. The cDNA sequence was analyzed and compared with the human cDNA library (hORFeomeV7.1). A list of proteins that interact with EspF was obtained and a protein interaction cloud plotted using the number of reads as references (Figure 2). We can see that CCDC75, DEGS1 and MRAP2 were significantly visible in the cloud as they possessed a relatively large number of reads. The number of reads is proportional to the intensity of the interaction between proteins. As is known, SNX9 has been demonstrated to interact with EspF [17]; we have indeed screened it, but it was not obvious in the cloud because the number of reads was not high. We suspect that the expression of this protein may be low, although an interaction may not be apparent because the number of reads is not high. Therefore, we thought that the reads were not a gold standard because the levels of protein expression will affect the reads.

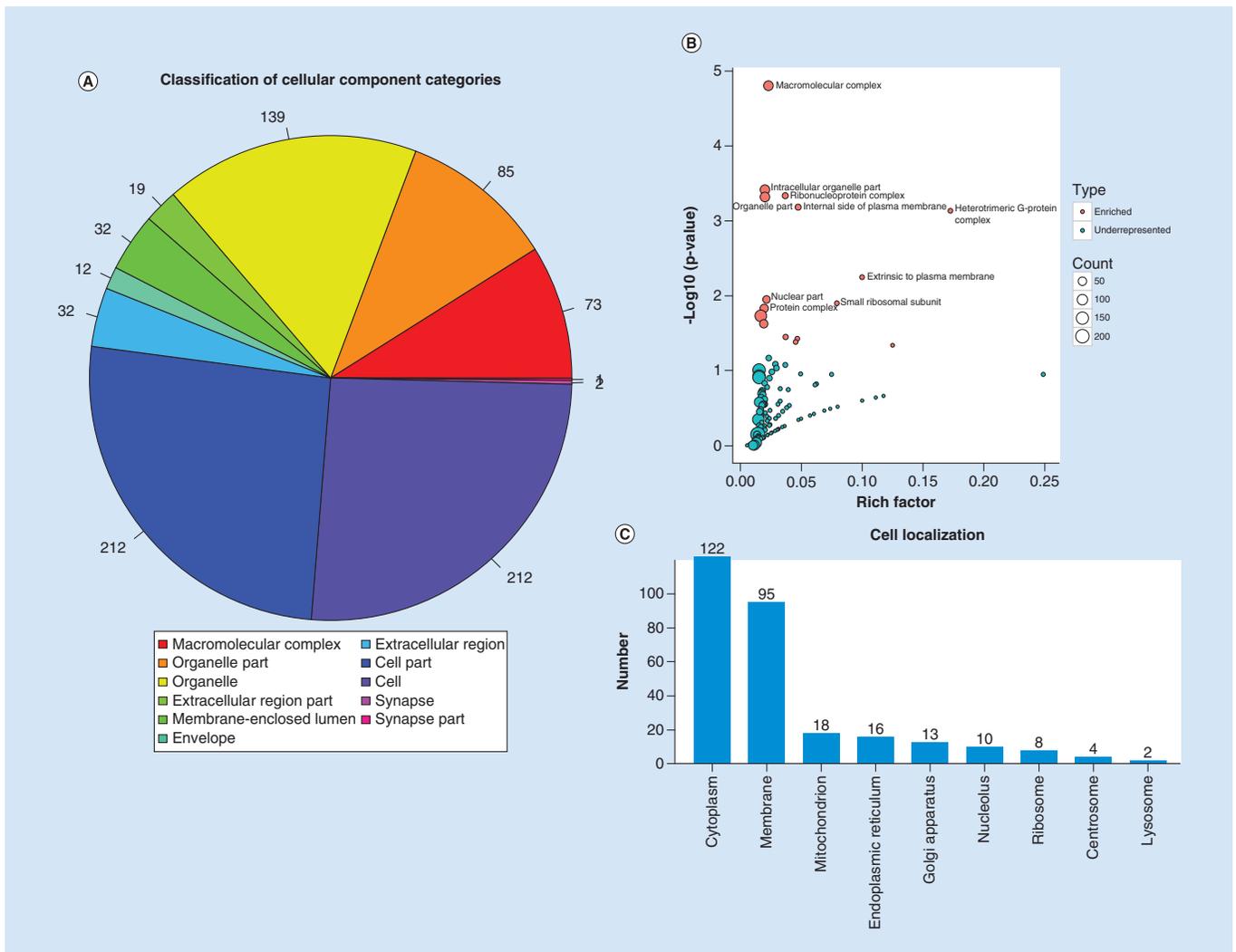


Figure 3. Analysis of cellular components and cell localization in enterohemorrhagic *Escherichia coli* O157:H7 EspF-interacting proteins. (A) GO annotation of identified EspF-related proteins in cellular components (CCs); proportions of gene CC terms are represented by pie charts; and the numbers at the edge of the pie charts indicate the number of genes associated with CC terms. (B) Distribution of enriched CCs – the horizontal axis represents the richness factor and the vertical axis represents p-value. Within each set analysis, a red circle represents a distinct significant functional CC cluster, and a blue circle represents an under-represented cluster. (C) Cell localization of EspF-related proteins.

Host proteins interacting with EHEC O157:H7 EspF were mostly macromolecular complexes & mainly located in the cytoplasm

The cellular composition and localization of the proteins were analyzed using the DAVID online analysis tool. A total of 212 proteins (72%) belonged to the cell fraction, 139 proteins (47%) belonged to organelles and 73 proteins (25%) belonged to macromolecule complexes (Figure 3). Target proteins were located in various cellular organelles such as mitochondria, ribosomes, endoplasmic reticulum and cytoplasm. Of these, 122 proteins (42%) were located in the cytoplasm, 95 proteins (33%) were located in the cell membrane and 18 proteins (6%) were located in mitochondria (Figure 3). Fisher’s exact test was used to enrich the data to exclude the influence of the background protein; 17 CC types were significantly enriched, the primary 3 of which were macromolecular complex ($p = 1.57 \times 10^{-5}$), intracellular organelle ($p = 0.0003$) and ribonucleoprotein complex ($p = 0.0004$; Table 1).

Table 1. Target genes mainly enriched in 17 cellular components[†].

GO ID	Cellular component	Proportion of genes (%) [‡]	p-value [§]	Genes
GO:0032991	Macromolecular complex	24.91	1.57E-05	<i>SOX2, EIF2C3, RPL23</i>
GO:0044446	Intracellular organelle part	29.01	0.0003	<i>SOX2, UQCR10, RPL23, PDHB</i>
GO:0030529	Ribonucleoprotein complex	6.48	0.0004	<i>EIF2C3, MRPS33, RPL22, STAU1</i>
GO:0044422	Organelle part	29.01	0.0004	<i>SOX2, UQCR10, RPL23, PDHB</i>
GO:0009898	Internal side of plasma membrane	4.77	0.0006	<i>GNG7, KRAS, ILK, RAB33A</i>
GO:0005834	Heterotrimeric G-protein complex	1.70	0.0007	<i>GNG10, GNG11, GNGT2, GNG7</i>
GO:0019897	Extrinsic to plasma membrane	1.70	0.0056	<i>GNG10, GNG11, GNGT2, GNG7</i>
GO:0044428	Nuclear part	13.31	0.0112	<i>SOX2, SMC1A, FTSJ3, RPL23</i>
GO:0015935	Small ribosomal subunit	1.70	0.0126	<i>MRPS33, RPS20, MRPS17, RPS12</i>
GO:0043234	Protein complex	17.40	0.0148	<i>SOX2, CALD1, FAM175A, PDHB</i>
GO:0005737	Cytoplasm	41.63	0.0185	<i>EIF2C3, ITM2A, UQCR10</i>
GO:0043232	Intracellular nonmembrane-bounded organelle	17.06	0.0235	<i>RPL23, CALD1, STAG3L4, MRPS17</i>
GO:0043228	Nonmembrane-bounded organelle	17.06	0.0235	<i>RPL23, CALD1, STAG3L4, MRPS17</i>
GO:0005840	Ribosome	2.73	0.0355	<i>MRPS33, RPL22, RPL23</i>
GO:0033279	Ribosomal subunit	2.04	0.0371	<i>MRPS33, RPL22, RPS20</i>
GO:0005681	Spliceosome	2.04	0.0415	<i>PPP1R8, SFRS12, SLU7, HNRNPK</i>
GO:0030532	Small nuclear ribonucleoprotein complex	1.02	0.0459	<i>LSM6, SLU7, SF3B3</i>

[†]In the enrichment analysis results, only cellular component classes with p-value < 0.05 are listed.

[‡]Percentage of target genes in the current class compared with annotated genes.

[§]Benjamini method used to control the error detection rate to carry out p-value correction.

EspF protein may mediate the inflammatory process by interacting with host proteins

The 293 proteins identified in this experiment were analyzed using the DAVID analysis system. These proteins were found to be involved in nearly 1000 BPs, such as cell physiology (58%), metabolic processes (39%) and bioregulatory processes (37%; Figure 4). Some of these proteins participated in many BPs, with enrichment analysis revealing that the target proteins were involved in 19 BPs (Table 2). The three primary enrichments were RNA splicing ($p = 0.005$), secreted protein regulation ($p = 0.012$) and positive regulation of acute inflammatory response ($p = 0.014$; Figure 4).

RNA splicing is a process to remove the noncoding sequences (introns) from mRNA precursors (pre-mRNAs), so that the coding sequences (exons) can be joined together to produce the mature mRNA. The mature mRNA was subsequently transported to the cytoplasm and translated into protein. This process contributes greatly to cell differentiation and development [35]. It has been demonstrated that EPEC infection can reduce the level of RPL9 protein in an EspF-dependent manner, which specifically threatens ribosome biosynthesis [36]. We speculated that the relationship between EspF and RNA splicing would eventually affect ribosomal proteins. Our bioinformatics analysis also found a close connection between EspF and ribosomes. Although the role of RNA splicing in the disease caused by EHEC infection is unclear, reduced ribosome biogenesis and protein translation may prevent defensive reactions during long-term bacterial infections.

EPEC and EHEC can cause A/E lesions and inject effectors by the T3SS, which are able to activate NF- κ B and MAPK pathways in response to their adherence to epithelial cells; this activation can promote proinflammatory responses. However, EPEC/EHEC have figured out ways to escape the inflammatory response and avoid the recruitment of immune cells, so as to avoid the elimination of pathogens [37]. It has been suggested that EPEC EspF inhibited the bacterial uptake in a PI3K-dependent manner to counteract macrophage phagocytosis, playing a crucial part in preventing bacterial uptake by macrophages [22]. EHEC infection can lead to hemorrhagic colitis and is the result of the combined effect of multiple proteins. Accumulating evidence suggests that EspF plays a key role in the pathogenesis of hemorrhagic colitis induced by *E. coli* O157:H7 by disrupting TJs between intestinal epithelial cells and disrupting mitochondrial function, leading to apoptosis [16,38]. We speculate that EspF plays

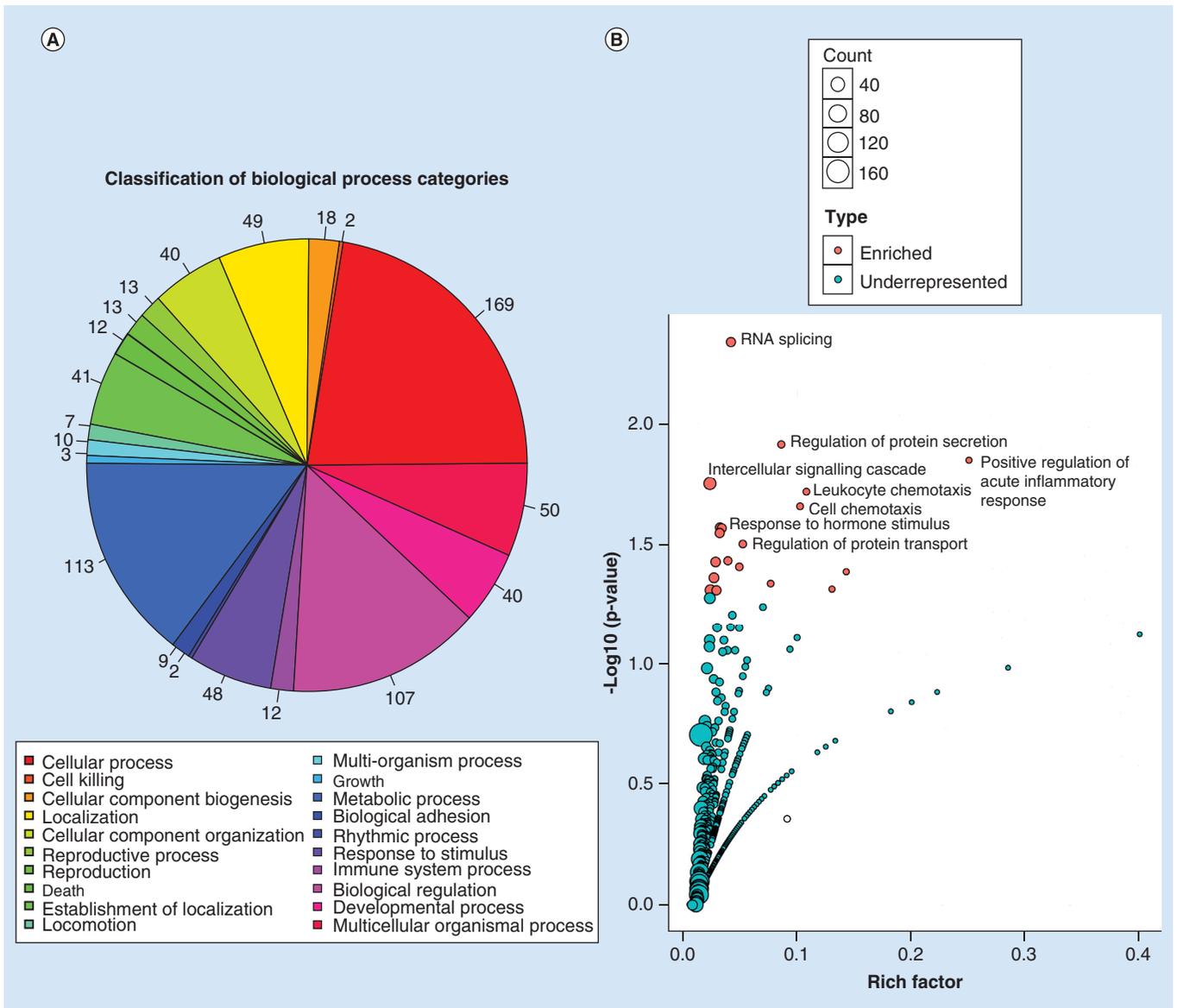


Figure 4. Analysis of biological process in enterohemorrhagic *Escherichia coli* O157:H7 EspF-interacting proteins. (A) GO annotation of identified EspF-related proteins in biological processes (BPs); proportions of gene BP terms are represented by pie chart, with the number of the genes associated with BP terms indicated outside the pie chart; **(B)** distribution of enriched BPs; horizontal axis represents the richness factor; vertical axis represents p-value. Within each set analysis, red circle represents a distinct significant functional cellular components cluster, the blue circle represents an under-represented cluster.

a similar double-edged sword role in the host, participating in, or even promoting, the host-cell inflammatory response; however, it can help to resist the host’s immune defense by inhibiting bacterial uptake, translocation and macrophage phagocytosis.

The results showed that EspF, as an important virulence protein, participates in the regulation of acute inflammatory response through interaction with host cells after injection to the host cell. It may participate in a number of important BPs, taking part in RNA splicing, regulating inflammation and even biological behavior such as tumor development. The specific mechanism remains unknown but is worthy of further study.

EspF showed GTPase activity, combined with splicing RNA & other potential molecular functions

We used the DAVID online tool to analyze MFs of the target proteins. These proteins were involved in a large number of MFs: 193 proteins (66%) had a binding activity, 77 proteins (26%) had catalytic activity and 18

Table 2. The target genes mainly enriched in 19 biological processes[†].

GO ID	Biological process	Proportion of genes (%) [‡]	p-value [§]	Genes
GO:0008380	RNA splicing	4.09	0.004	<i>SMC1A, PPP1R8, SRPK1, PPP2CA</i>
GO:0050708	Regulation of protein secretion	1.70	0.012	<i>CD40LG, CRTAM, HDX, SAA1</i>
GO:0002675	Positive regulation of acute inflammatory response	1.02	0.014	<i>CCL5, TAC1, IL6</i>
GO:0007242	Intracellular signaling cascade	10.23	0.017	<i>SMC1A, PRDX4, GNG7, KRAS</i>
GO:0030595	Leukocyte chemotaxis	1.36	0.018	<i>CCL5, SAA1, SBDS, IL6</i>
GO:0060326	Cell chemotaxis	1.36	0.021	<i>CCL5, SAA1, SBDS, IL6</i>
GO:0009725	Response to hormone stimulus	4.09	0.026	<i>GNG10, GNG11, CCL5, TFF3</i>
GO:0006397	mRNA processing	3.75	0.026	<i>SMC1A, PPP1R8, SRPK1, SFRS12</i>
GO:0016071	mRNA metabolic process	4.09	0.028	<i>EIF2C3, SMC1A, PPP1R8, SRPK1</i>
GO:0051223	Regulation of protein transport	2.04	0.031	<i>CD40LG, CRTAM, HDX, PKIG</i>
GO:0051046	Regulation of secretion	2.73	0.036	<i>CD40LG, CRTAM, HDX, TAC1</i>
GO:0051276	Chromosome organization	4.77	0.037	<i>SOX2, SMC1A, H3F3B</i>
GO:0070201	Regulation of establishment of protein localization	2.04	0.038	<i>CD40LG, CRTAM, HDX</i>
GO:0002673	Regulation of acute inflammatory response	1.02	0.040	<i>CCL5, TAC1, IL6</i>
GO:0006396	RNA processing	5.11	0.043	<i>SMC1A, FTSJ3, SRRT</i>
GO:0009755	Hormone-mediated signaling	1.36	0.045	<i>GNG10, GNG11, GNG7, GNG12</i>
GO:0000096	Sulfur amino acid metabolic process	1.02	0.048	<i>MTHFD2L, MAT2A, NFS1</i>
GO:0051716	Cellular response to stimulus	6.82	0.048	<i>SMC1A, SRRT, TOPBP1</i>
GO:0009719	Response to endogenous stimulus	4.09	0.048	<i>GNG10, GNG11, CCL5</i>

[†] Biological process classes with p-value < 0.05 are listed in the enrichment analysis results.

[‡] Percentage of target genes in the current class compared with annotated genes.

[§] Benjamini method used to control the error detection rate to carry out p-value correction.

Table 3. The target genes mainly enriched in six molecular functions[†].

GO ID	Molecular function	Proportion of genes (%) [‡]	p-value [§]	Genes
GO:0003735	Structural constituent of ribosome	2.73	0.012	<i>MRPS33, RPL22, RPL23</i>
GO:0003924	GTPase activity	3.07	0.013	<i>GNG10, GNG11, RAB30, GNGT2</i>
GO:0005515	Protein binding	46.41	0.030	<i>EIF2C3, ITM2A, ZCCHC3, RPL23</i>
GO:0003723	RNA binding	6.14	0.038	<i>EIF2C3, RPL22, STAU1, KHDC1</i>
GO:0031202	RNA splicing factor activity/transesterification mechanism	1.02	0.048	<i>SLU7, SF3B3, TXNL4A</i>

[†] Molecular function classes with p-value < 0.05 are listed in the enrichment analysis results.

[‡] Percentage of target genes in the current class compared with annotated genes.

[§] Benjamini method used to control the error detection rate to carry out p-value correction.

proteins (6%) had molecular transport activity (Figure 5). Enrichment analysis showed that the target proteins were enriched in five main types of MF: ribosome structure (p = 0.012), GTPase activity (p = 0.013), protein binding (p = 0.030), RNA binding (p = 0.038) and RNA splicing activity (p = 0.048; Table 3). Studies have shown that EspF has the ability to target mitochondria and nucleoli [6,16], and interacts with N-WASP, SNX9 and Abcf2 [8,17–18], leading to microvilli disappearance, cytoskeleton rearrangement, actin aggregation, mitochondrial dysfunction and apoptosis in intestinal epithelial cells, nucleolysis and other functions [11,39]. EspF with GTPase activity combined with splicing RNA and other functions in the host cell remains unknown research territory, deserving of further investigation.

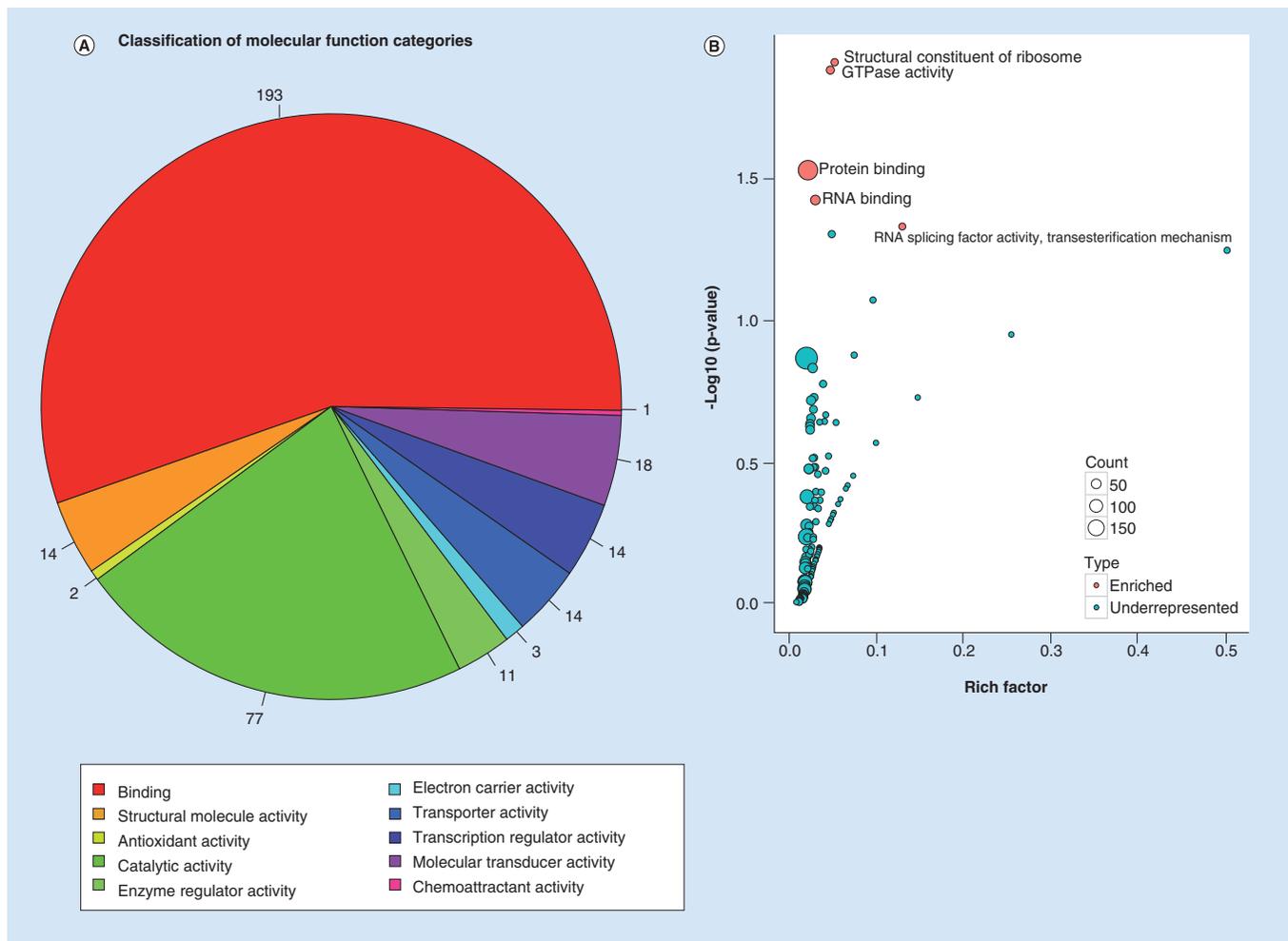


Figure 5. Analysis of molecular functions in enterohemorrhagic *Escherichia coli* O157: H7 EspF-interacting proteins. (A) GO annotation of identified EspF-related proteins in molecular function (MF); proportions of gene MF terms are represented by pie chart, with numbers of genes associated with MF terms shown at the edge. **(B)** Distribution of enriched MFs – horizontal axis represents the richness factor, vertical axis represents p-value. Within each set analysis, red circle represents a distinct significant functional cellular components cluster, blue circle shows an under-represented cluster.

EspF may play its pathogenic role by regulating the splice & ribosomal signal pathways

DAVID was used for KEGG signaling pathway analysis, finding that the target proteins were involved in 109 signaling pathway types. Of these, nine proteins including GNG10, CCL19 and CCL5 (3%) were involved in chemokine signaling pathways, seven proteins including PIP4K2B, MYL10 and ENPP4 (2.4%) were involved in regulation of actin cytoskeleton signaling pathways, seven proteins including KRAS, FGF7 and LAMB1 (2.4%) were involved in cancer signaling pathways and six proteins including DUSP14, PPM1B and GNG12 (2%) were involved in MAPK signaling pathways (Figure 6). There were three types of signaling pathway with significant enrichment: chemokine signaling ($p = 0.005$), splicing signaling ($p = 0.033$), and ribosome signaling pathways ($p = 0.035$; Table 4). Strikingly, the sixth pathway: regulation of the actin cytoskeleton, and the eighth pathway: TJ, aroused our concern. The breakdown of TJs allowed the lumen material to penetrate the epithelial cell intercellular spaces, which can cause inflammation and then lead to diarrhea [40]. It has been reported that both EPEC and EHEC can induce rearrangements of the actin cytoskeleton [41]. The EPEC EspF protein collaborated with the effectors Map, Tir, EspG and NleA to interfere with TJs [42–44]. Moreover, EspF can cause actin sequestration and recruit junctional proteins, such as occludin, claudin, ZO-1 and ZO-2, to the pedestals, inducing the redistribution and loss of transepithelial electrical resistance, disruption of paracellular permeability and depolymerization of actin. Ultimately, this leads to EspF-induced TJ disruption [45]. Our study suggests that EspF may be involved in

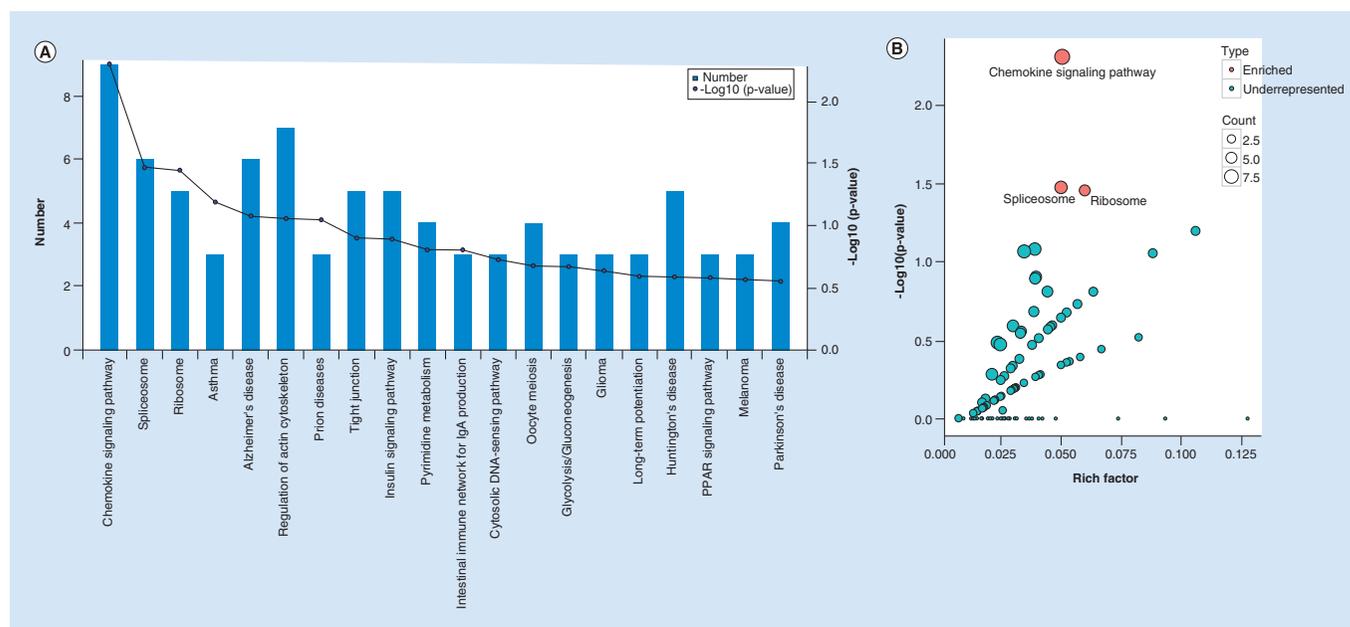


Figure 6. Analysis of the Kyoto Encyclopedia of Genes and Genomes pathway in enterohemorrhagic *Escherichia coli* O157:H7 EspF-interacting proteins. (A) The distribution of the top 20 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Horizontal axis shows related pathways; the left vertical axis shows the number of pathways; the right vertical axis shows the p-value. The blue line shows the signal pathway enrichment distribution trend. **(B)** KEGG pathway enrichment bubble chart. Horizontal axis represents the rich factor; vertical axis represents p-value. Red circles represent a distinct significant KEGG pathway cluster; blue circles represent an under-represented cluster.

the regulation of the actin cytoskeleton and TJs, which confirmed and extended previous findings that indicated the disruption of the TJ barriers might be associated with EspF's regulation of the actin cytoskeleton.

Analysis of BPs and MFs showed that EspF interacted with ribosomal protein, and the KEGG signaling pathway enrichment results further complemented the relationship between EspF and ribosome. These three enriched signaling pathways play an important role in cell growth, development and differentiation, as well as apoptosis, indicating that EspF is involved in many aspects of human life activity and in disease processes. In particular, its interaction with ribosomal protein is worthy of further study.

EspF may induce colitis & colorectal carcinogenesis through a major protein–protein interaction network

We analyzed the interaction of EspF with the target proteins by STRING online tools, focusing on the significantly enriched proteins in BPs, MFs and KEGG pathways and using cytoscape to map the protein interaction network. It was found that 89 and 11% of the listed proteins had multidirectional and unidirectional interaction networks, respectively. Within the protein interaction network, RPS, the RPL protein family, and KRAS protein had a higher connective degree (Figure 7). RPS and RPL proteins are types of ribosomal protein. Research has shown that ribosomal proteins play a pivotal role in cell proliferation, differentiation and apoptosis, in addition to protein synthesis and the transcriptional regulation of genes [46,47]. The RAS family plays an important regulatory role in the processes of cell growth and differentiation, with KRAS acting as a molecular switch and participating in a number of cancer pathways such as that of colorectal cancer [48–50]. Studies have shown that 40–50% of human colorectal cancer is caused by the activity of KRAS protein expression, and inhibition of KRAS expression can inhibit tumor growth [51–53]. In recent years, a hypothesis that adherent A/E *E. coli* has a connection with colorectal carcinogenesis in humans has been put forth [54–56]. It has been shown that EPEC infection can cause phenomena such as dramatically enhancing the spontaneous mutation frequency in host cells, inducing mismatch repair (MMR) protein dysfunction and increasing the levels of host reactive oxygen species, which were considered to be symbols of tumorigenesis. These were dependent on EspF's mitochondrial targeting [57]. Although there is no direct evidence for the relationship between infection and colorectal carcinogenesis, our results provide us with

Table 4. The top 20 Kyoto Encyclopedia of Genes and Genomes pathways of the target genes[†].

MAP ID	Kyoto Encyclopedia of Genes and Genomes pathway	Proportion of genes (%) [‡]	p-value [§]	Genes
hsa04062	Chemokine signaling pathway	3.07	0.004	<i>GNG10, GNG11, CCL5, KRAS</i>
hsa03040	Spliceosome	2.04	0.033	<i>LSM6, SLU7, HNRNPK, SF3B3</i>
hsa03010	Ribosome	1.70	0.035	<i>RPL22, RPL23, RPS20, RPS12</i>
hsa05310	Asthma	1.02	0.064	<i>CD40LG, HLA-DQA1, IL9</i>
hsa05010	Alzheimer's disease	2.05	0.083	<i>COX5A, UQCR10, NDUFB1, CALML5</i>
hsa04810	Regulation of actin cytoskeleton	2.39	0.086	<i>PIP4K2B, MYL10, ENPP4, KRAS</i>
hsa05020	Prion diseases	1.02	0.089	<i>CCL5, IL6, PRNP</i>
hsa04530	Tight junction	1.71	0.125	<i>MYL10, PPP2CA, PRKCH, KRAS</i>
hsa04910	Insulin signaling pathway	1.71	0.127	<i>HK1, PCK2, CALML5, EIF4E</i>
hsa00240	Pyrimidine metabolism	1.37	0.155	<i>CMPK1, UCK2, NME5, POLR1D</i>
hsa04672	Intestinal immune network for IgA production	1.02	0.155	<i>CD40LG, HLA-DQA1, IL6</i>
hsa04623	Cytosolic DNA-sensing pathway	1.02	0.186	<i>CCL5, IL6, POLR1D</i>
hsa04114	Oocyte meiosis	1.37	0.209	<i>SMC1A, CALML5, PPP2CA, CHP</i>
hsa00010	Glycolysis/gluconeogenesis	1.02	0.212	<i>HK1, PCK2, PDHB</i>
hsa05214	Glioma	1.02	0.228	<i>CALML5, KRAS, E2F2</i>
hsa04720	Long-term potentiation	1.02	0.255	<i>CALML5, KRAS, CHP</i>
hsa05016	Huntington's disease	1.71	0.258	<i>AP2A2, COX5A, UQCR10, NDUFB1</i>
hsa03320	PPAR signaling pathway	1.02	0.260	<i>SCD5, PCK2, ILK</i>
hsa05218	Melanoma	1.02	0.271	<i>KRAS, FGF7, E2F2</i>
hsa05012	Parkinson's disease	1.37	0.278	<i>COX5A, UQCR10, NDUFB1, NDUFA7</i>

[†]The top 20 Kyoto Encyclopedia of Genes and Genomes pathways of the target genes are listed; the first three are Kyoto Encyclopedia of Genes and Genomes pathway classes with p-value < 0.05.
[‡]Percentage of target genes in the current class compared with annotated genes.
[§]Benjamini method used to control the error detection rate to carry out p-value correction.

a map of candidate targets. Exploring the interaction of EspF with these proteins will help to further elucidate the pathogenic mechanism of EspF in colitis and colorectal carcinogenesis.

The plasmid-encoded native EspF is subcellularly localized to mitochondria in a manner dependent on its N-terminal domain

We cloned the *espF* gene and its N-terminal and C-terminal sequences into the eukaryotic expression vector pEGFP-N1 to express EspF, EspF/N and EspF/C *in vitro*, respectively, and observed their localization after the transfection of HeLa cells. The results showed that the EspF and EspF/N proteins were localized in mitochondria and showed a spotted, filamentous distribution (Figure 8), which is consistent with the results of the Dean's study [36]. The plasmid encoding the EspF protein showed a dotted distribution, and it was reported previously that EspF may be a soluble cytoplasmic protein with a nonhomogeneous distribution [38]. Interestingly, we found that the EspF/C protein was homogeneously distributed (Figure 8), which is completely different from the EspF and EspF/N proteins, and it appears to be distributed throughout the entire cell and not specifically located in any organelle. The way the protein is expressed may affect their function in host cells, we suspect that the N-terminal is more important determining the targeting of the EspF protein, and the C-terminal protein may determine other functions of the EspF protein. Dean and Kenny have proven that ectopic expression of EPEC EspF induced multinucleation and cell–cell internalization in small intestinal epithelial cells, which were dependent on an unknown protein–protein interaction site in its C-terminal repeat region [58]. Our results further demonstrate that the EspF protein still has the function of targeting the mitochondria *in vitro*, and that the N-terminal is the decisive factor in the targeting of mitochondria.

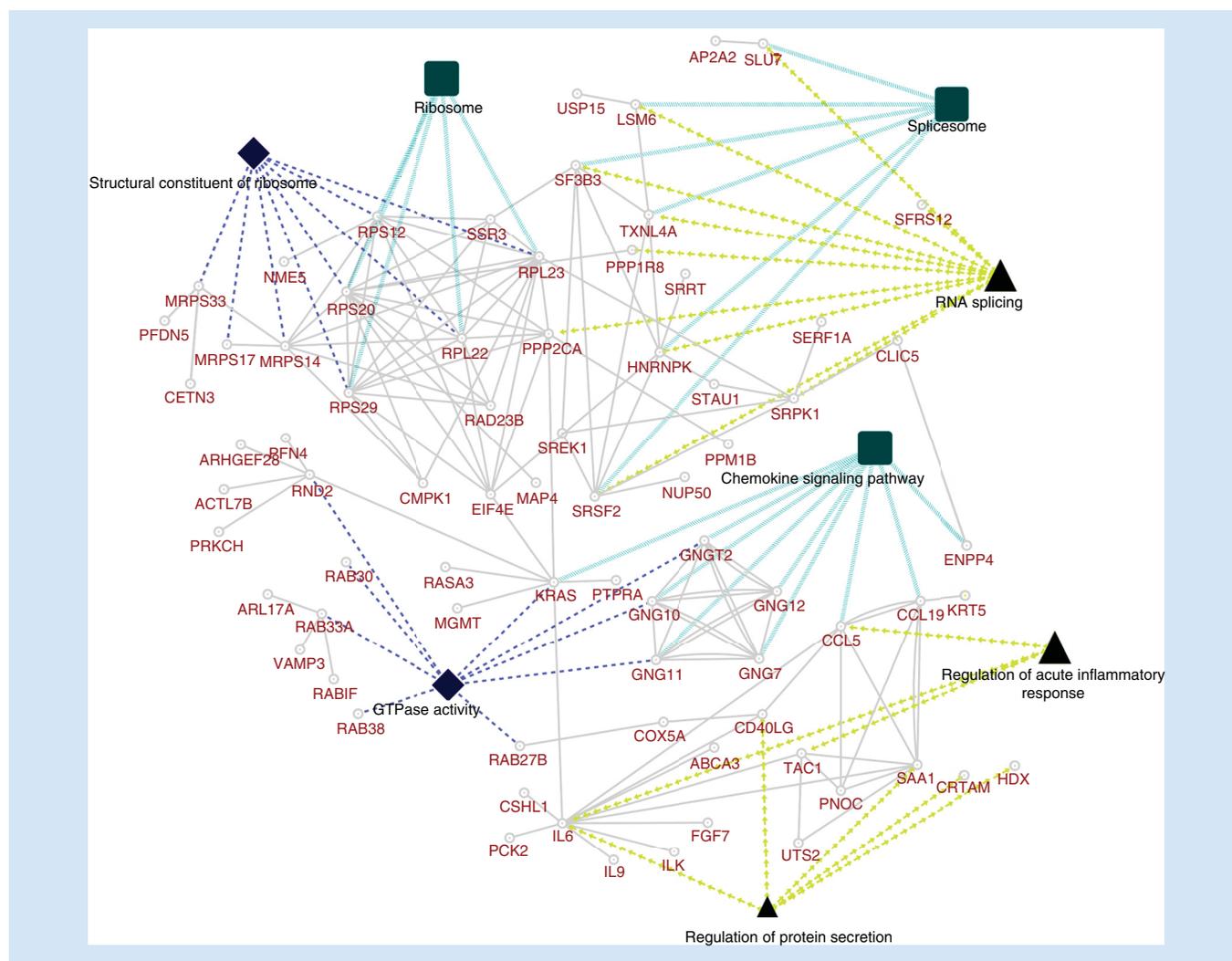


Figure 7. Analysis of statistically significant pathways detected in enterohemorrhagic *Escherichia coli* O157:H7 EspF-interacting proteins. Statistically significant biological processes in the network are mapped as triangular nodes; molecular functions are represented by diamonds; and Kyoto Encyclopedia of Genes and Genomes pathways by squares. Color codes show clustering by functional similarity with ungrouped functional outliers shown in gray. Small nodes represent the proteins, labeled with red gene ID names, are associated with the statistically significant GO and Kyoto Encyclopedia of Genes and Genomes terms. The networks also illustrate the functional relationships (the edges) between the nodes, the thickness of which is directly proportional to the association significance score.

EspF binds to ANXA6 *in vitro*, which may result in the inhibition of phagocytosis

To confirm the screening results, we chose 12 proteins and tagged them with the FLAG tag. The EspF protein was tagged with GST, then coexpressed GST-tagged EspF and FLAG-tagged protein in 293T cells, respectively, and a GST pull down and dot blotting were used to verify the interactions of these proteins. The results showed that SNX9 and ANXA6 had a strong positive reaction (Figure 9). It was reported that the EspFs of both EHEC and EPEC had the ability to bind to SNX9, but EHEC EspF had weaker binding to SNX9 compared with EPEC [59]; our results confirmed and extended the previous results that EHEC EspF can still bind to SNX9 when it is expressed *in vitro*.

ANXA6 emerged as a multifunctional scaffold protein, and its ability to recruit signaling proteins, interact with actin and organize the cytoskeleton, and modulate membrane trafficking allowed it to be involved in many biological processes, such as membrane repair, cell proliferation, differentiation, survival and inflammation [60]. It has been suggested that ANXA6 cooperates with PKC α to downregulate EGFR signaling through elevation of PKC α and EGFR/PKC α membrane-targeting levels [61]. Strikingly, EspF can mediate EGFR cleavage and loss dependent on caspase activation at the late stages of EPEC infection [62]. The levels of EGFR represent the rate

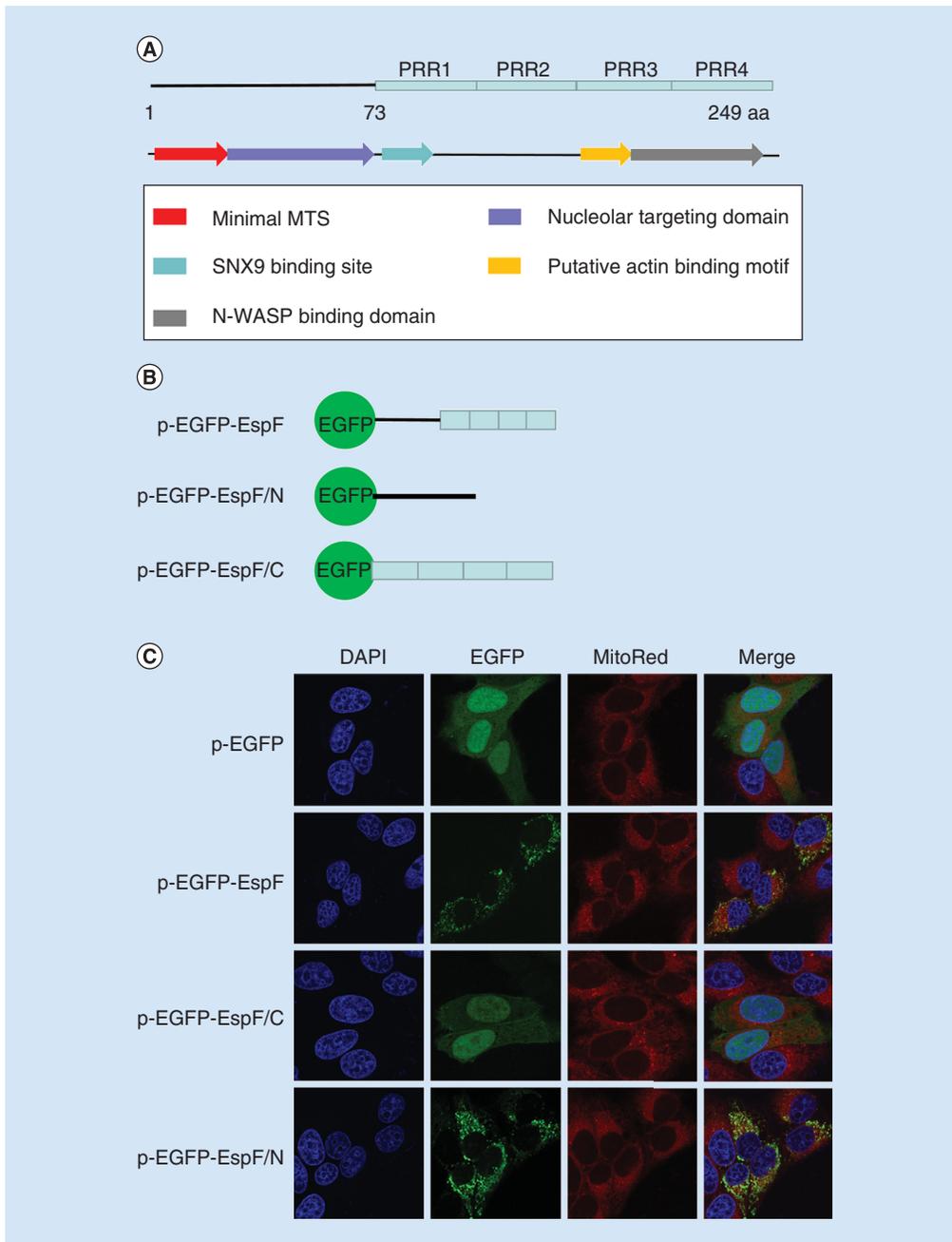


Figure 8. The mitochondrial localization of EspF is dependent on its N-terminal domain. (A) The domain of the enterohemorrhagic *Escherichia coli* espF gene. **(B)** The general structures of the eukaryotic expression vectors pEGFP-EspF, pEGFP-EspF/N and pEGFP-EspF/C. **(C)** Fluorescence images of EspF subcellular localization were captured using an EGFP-fused eukaryotic expression vector transfected in HeLa cells. Cell nuclei were stained with DAPI (blue), and cell mitochondria were stained with MitoTracker (red). EspF/N protein was colocalized with the mitochondria, as well as the EspF protein, and these proteins were dotted, with a filamentous distribution. The EspF/C protein was distributed in the cytoplasm and nuclei, as well as the EGFP protein, and they were homogeneously distributed.

of cell endocytosis, degradation, recycling and synthesis [63]. Although the function of EGFR in EHEC infection has not been studied, the interaction between EspF and ANXA6 could act on EGFR and be one mechanism that contributes to the inhibition of phagocytosis by macrophages. Also, we cannot rule out the possibility that EspF collaborates with ANXA6 to rearrange the actin cytoskeleton. This provides us with a new view, and this deserves further research.

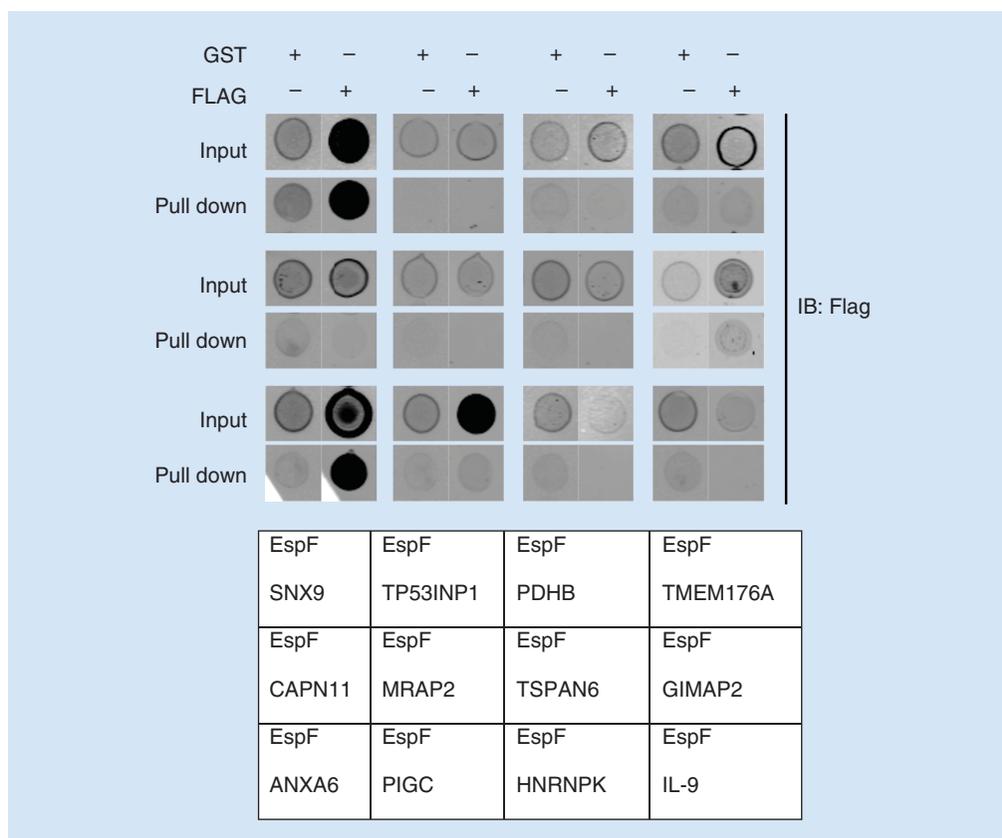


Figure 9. Co-precipitation screens to identify the protein–protein interactions between EspF and prey proteins. The selected 12 candidate proteins were tagged with FLAG, the EspF protein was tagged with glutathione S-transferase, and were cotransfected into 293T cells. Cell proteins were extracted and incubated with GSH–agarose beads. Then, the cell proteins (input) and the GSH beads binding the proteins (pull down) were dot blotted and probed with anti-glutathione S-transferase–HRP or anti-FLAG–HRP antibodies.

The reads of ANXA6 were low, but coprecipitation results showed a strong positive signal; therefore, we hypothesize that the expression of ANXA6 protein in the BiFC system was poor, which may have directly resulted in low protein reads. Meanwhile, coprecipitation results revealed that PIGC, TMEM176A and GIMAP2 have a moderate positive reaction, and CAPN11 and PDHNB have a weak positive reaction (Figure 9). The correlation between these proteins and EspF deserve our close attention as well. TP53INP1, MRAP2, TSPAN6, HNRNPK and IL-9 failed to bind EspF in the coprecipitation test (Figure 9); they may be false positives, or require different conditions or intermediate factors for coprecipitation to occur.

Discussion

Since most of the data related to EspF are from research on EPEC rather than EHEC, we would like to focus on the EHEC EspF protein. The importance of EspF as a versatile effector cannot be overstated, and its interaction with host proteins is particularly significant in the pathogenic mechanisms of EHEC O157:H7. The confirmed EspF-interacting proteins are mostly obtained by yeast-two-hybrid methods, but whether the interaction occurs in mammals is not known. In our study, we employed a BiFC method that takes advantage of the split Venus YFP to examine the interplay between EspF and human proteins.

Studies have demonstrated the advantage of using BiFC to detect protein–protein interactions [64,65]. We noticed that the BiFC technology has at least three advantages that are not found with conventional approaches such as yeast-two-hybrid assays or co-immunoprecipitation (co-IP)/mass spectrometry. First, the fluorescence not only indicates protein–protein interactions and the interaction location but also permits other techniques, such as flow cytometry, to be used to sort the positive signals [66]; second, the presence of fluorescence not only means the formation of protein complexes but also means the formation of strong hydrogen bonds between the N- and

C-termini of YFP, thus, BiFC can detect transient and weak interactions [67]; and third, BiFC is the optimal method for single or live-cell research because co-IP/mass spectrometry needs a large number of cells and lengthy purification steps *in vitro*, which may prevent the detection of low affinity interactions.

In a previous genome-wide BiFC screen done by Songyang *et al.*, the same system as ours was used, with BiFC screening with a smaller YFPc library of 12,000 genes. The researchers found 300 positive proteins from BiFC assays. They then tagged the 300 genes with FLAG tag and coexpressed them individually with their respective GST-tagged bait proteins, determining results by GST pull down. They found that about 72% of the interactions found using BiFC could be verified by GST pull down [27,68]. We conducted largescale, high-precision screening of the proteins interacting with EHEC O157:H7 EspF, using a BiFC platform and flow-sorting technology for the first time. The result was a list of 293 proteins, and we analyzed the BPs, MFs, cellular localization and signaling pathways involved in these proteins using bioinformatic methods.

Previous studies have shown that EspF can interact with some proteins. Jean-Philippe applied a yeast-two-hybrid system and found that EPEC effector EspF interacts with the host protein Abcf2 [8]; Oliver Marche's identified that SNX9 can bind EPEC EspF via its amino-terminal SH3 region [17]. Both EHEC and EPEC EspF possess 47 aa PRRs; Alto provided further proof that EspF PRR1 binds the SH3 domain of SNX9 and binds the CRIB domain of N-WASP [39]. In our results, we found that SNX9, SNX18 and EHEC O157:H7 EspF proteins also interact. In the dot-blotting experiments, in spite of low protein reads, SNX9 showed a strong positive signal, which indicates it has strong bonding with EspF. Unfortunately, we did not screen for Abcf2 and N-WASP, but screened for the same ATP-binding cassette family of ABCA3. First, the *espF* gene sequences of EPEC and EHEC were 87% identical, although they had high similarity, their pathogenicity appeared to be different, for example, in reducing the function of epithelial resistance (TER) [69]. Second, the N-terminal of EspF protein in EPEC and EHEC includes the bacterial secretion signal site and the two organellar targeting domains. However, a number of PRR repeats were different: EPEC has three PRRs (PRR1–PRR3), EHEC has four PRRs (PRR1–PRR4) and the SNX9 and N-WASP binding site is located at PRR1 [11]. The functions of the PRR domains are different; meanwhile, considering the N-terminal is a functional domain, we set the YFPn tag at PRR4 of the EspF protein when designing the experiment. The YFPn tag may alter the conformation or activity of the labeled protein, having an effect on the binding site on PRR1. Besides, the combination between N-WASP and EspF can be detected by coimmunoprecipitation at 1–1.5-h postinfection, but no interplay was found at 3-h postinfection [45]; this may also be one of the reasons why we did not screen N-WASP. It can be seen that the protein reads were not always proportional to the interaction strength, but it was influenced by the expression level of the protein itself; therefore, we set it as a preliminary criterion for protein interaction, not the gold standard.

Interestingly, in our study we found AQP7P2, a type of water channel protein. AQP7 is thought to be the sole adipose glycerol channel [70], and can maintain basic life activities by transporting tens of thousands of water molecules through an 'hourglass' model [71]. For the first time, we found aquaporins in the study of EHEC, which may interact with EspF, opening a new door for the further study of EHEC-induced diarrhea.

Consistent with the previous study, we also found a profilin protein (PFN4), which is a small actin-binding protein, in our protein cloud. By interacting with numerous proteins such as actin, WASP, GTPase and annexin 1, profilins can regulate actin polymerization or other cellular processes, such as membrane trafficking, small-GTPase signaling and nuclear activities [72]. Profilin has a vital role in the regulation of actin cytoskeleton and by cooperating with actin and EspF, they form a complex which enables EspF to serve as a nucleation-promoting factor [45]. KEGG signaling pathway analysis results indicated that EspF may disrupt TJs through manipulation of the actin cytoskeleton; we hypothesize that PFN4 may play a role in this process by interacting with EspF, and the interplay may affect the polymerization of actin, disrupting the formation of pedestals and leading to the redistribution of TJ proteins. Moreover, the interplay may possibly perform different functions, such as the promotion of nucleation.

In our second screen by coprecipitation, a protein that acts similarly to EspF–ANXA6 was found to interact with EspF. The interaction with actin gave ANXA6 the capacity to rearrange the cytoskeleton and form membrane–cytoskeleton complexes [60]. ANXA6 may recruit complex protein–protein interactions to remodel the actin cytoskeleton, which is the vital middle chain for EspF's ability to destroy TJs. Besides, increasing evidence has highlighted the actin cytoskeleton's role in regulating EGFR recycling and controlling EGFR endocytosis and degradation [73–75]. ANXA6 can also interact with the actin cytoskeleton and may guide endocytic vesicles to multivesicular bodies/late endosomes. Meanwhile, ANXA6 can negatively control the EGFR/Ras pathway through binding to p120GAP and PKC α [76]. Both EPEC and EHEC infections can modulate the host cytoskeleton by

activating PKC α and recruiting PKC to form adhesion pedestals via functionally intact lipid rafts [77–79], but the precise factor that mediates this event is unknown.

PKC activation is also controlled by the PI3K/Akt signaling pathway, on which the EPEC depends to escape phagocytosis by host cell macrophages [78,79]. EPEC has been reported to engage and activate EGFR signaling to cause EGFR autophosphorylation [80]. Although the phosphorylated EGFR increased the survival rate of host cells in early infection, EspF facilitated EGFR loss in the later infection process, leading to a sharp increase in host cell death [62]. The detection of ANXA6 bonding to EspF could provide a new possible mechanism for EspF; we assume that once EHEC infects host cells, PKC α will be activated, and EspF will be injected and interact with ANXA6 or form a complex with ANXA6 and actin, downregulating EGFR through the regulation of the actin cytoskeleton and manipulating the process of phagocytosis to help EHEC survive within macrophages.

In agreement with previous studies, we also found many signaling pathways in which EHEC is involved, including the Toll-like receptor signaling pathway, the actin skeleton regulatory pathway and the MAPK signaling pathway. The KRAS protein was involved in 34 signaling pathways, playing a leading role in the protein network. The activation of KRAS is essential for the development and progression of colorectal cancer [81], and the continuous accumulation of mutations of *kras* gene can drive healthy colon epithelial cells to develop dysplastic adenomas and even lead to colorectal cancer [82,83]. Mutations in KRAS, BRAF and PIK3CA (PI3K) were described as the types most related to colorectal cancer, and pathogenic *E. coli*, such as EPEC, could be a cofactor in the pathogenesis of colorectal cancer [56]. EPEC-induced cancer cell detachment and survival relies on the induction of MIC-1, implicating EPEC as a potent cancer trigger [84]. A cohort study showed that *E. coli* adherence to colonic mucosa may be linked to colorectal cancer [85]. Moreover, EPEC infection can activate PKC α , phosphorylate Bcl-2 and inhibit the transcription of MMR genes [86]. A/E *E. coli* infected intestinal epithelial cells by injecting toxic proteins, leading to DNA damage in host cells, and, together with the Toll-like receptor signaling pathway, the NF- κ B pathway and other cell inflammatory pathways, accelerated the process of cancer [86–88]. A/E *E. coli* can regulate the expression of DNA mismatch repair protein in host cells by post-transcriptional manipulation of EspF protein, causing a decrease in host cell MMR protein levels, EspF entered the host cell, targeting binding to mitochondria, leading to Apc and MMR protein gene mutations [57], and may even activate *kras* and other oncogenes, leading to changes in intestinal epithelium, causing colitis and even colorectal carcinogenesis. The occurrence of carcinogenesis is a multifactorial outcome, and it is a complex event caused by chronic inflammation and mutagenesis [84]. EspF may employ other host proteins such as ANXA6, KRAS, PKC α and other proteins to play a pathogenic role. The interaction between the EspF and KRAS proteins provides an important opportunity for the study of hemorrhagic colitis and even colorectal carcinogenesis during a long period of EHEC infection, and we will conduct further research.

We have verified that other proteins interact with EspF in the coprecipitation assay, including PIGC, TMEM176A, GIMAP2, CAPN11 and PDHB. Although there are no relevant studies that show the mechanisms of their interaction in EHEC infection, it was inspiring to determine to what extent these data will be translated to EHEC pathogenicity in the future. Studies are underway to further reveal the particular mechanism.

Undeniably, there are some limitations of the BiFC method in the screening process. One is that target proteins must be tethered to the two fragments of YFP and be coexpressed in the same cells. YFP tags may affect the conformation or activity of the tagged proteins, thus leading to false negatives. Second, some interactions may be dependent on cell-type or regulated developmentally; moreover, the delayed fluorescent signal and the irreversible BiFC process may affect the dissociation and association of protein complexes. Third, bait and prey proteins expression levels in the cells were out of control, and high expression levels may lead to false positives. Further research is needed to unravel how these proteins contribute to EspF's role in the pathogenicity of EHEC infection.

Our research results suggested that the 293 identified proteins may not bind simultaneously with the EspF protein. They may interact with EspF only or form a variety of subcomplexes and associate with EspF; it is equally possible that their interactions depend on time, space and cell type. Confirmation of the interaction of ANXA6 and EspF provides us with a new insight into EspF's function in the regulation of EGFR and antiphagocytosis. Further studies are needed to unravel how these proteins interact with EspF, what process they contribute to, and how to mediate bacteria–host pathogenesis. The interaction networks between EspF and host protein in human cells may be much more complex than we assumed. Our screenings offer a multitude of candidates for the further analysis of their potential role in EspF biology and EspF-induced signaling pathways, laying a solid foundation for the further study of the pathogenesis of EHEC.

Conclusion & future perspective

In this study, we focused on the host proteins that interact with EHEC O157:H7 EspF and handle the multiple aspects of the screened 293 host proteins. These findings provided specific candidates that interact with EspF and showed that EspF has the ability to interact with featured ribosomal proteins and to modulate the antiphagocytosis process by bonding with ANXA6. These results regarding the interaction between EspF and human proteins are very inspiring and provided a high-resolution map of the EspF interactome, which holds great promise for the area of protein interactions between the pathogen EHEC O157:H7 and host cells.

Summary points

- We screened 293 host proteins interacting with enterohemorrhagic *Escherichia coli* EspF using a bimolecular fluorescence complementation method.
- EspF and its N-terminal protein were colocalized with mitochondria when expressed *in vitro*.
- EspF's participant in regulating actin cytoskeleton and tight junctions was suggested, indicating that EspF may disrupt tight junctions through the modulation of actin cytoskeleton.
- The interactions between ANXA6 and EspF were verified; the combination of them may mediate actin cytoskeleton rearrangement, downregulate EGFR signal to develop antiphagocytosis function.
- EspF may employ a large protein interaction network to play a distinctive pathogenic role in enterohemorrhagic *E. coli* infection process.

Financial & competing interest's disclosure

This work was supported by the National Natural Science Foundation of China (No. 81371765) and the Natural Science Foundation of Guangdong Province (No. 2014A030313312). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

The authors thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Zhao S, Zhou Y, Wang C *et al*. The N-terminal domain of EspF induces host cell apoptosis after infection with enterohaemorrhagic *Escherichia coli* O157:H7. *PLoS ONE* 8(1), e51164 (2013).
2. Rangel JM, Sparling PH, Crowe C, Griffin PM, Swerdlow DL. Epidemiology of *Escherichia coli* O157:H7 outbreaks, United States, 1982–2002. *Emerg. Infect. Dis.* 11(4), 603–609 (2005).
3. Kaper JB, Nataro JP, Mobley HLT. Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* 2(2), 123–140 (2004).
4. Riley LW, Remis RS, Helgerson SD *et al*. Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *New Engl. J. Med.* 308(12), 681–685 (1983).
5. Garmendia J, Frankel G, Crepin VF. Enteropathogenic and enterohemorrhagic *Escherichia coli* infections: translocation, translocation, translocation. *Infect. Immun.* 73(5), 2573–2585 (2005).
6. Dean P, Kenny B. The effector repertoire of enteropathogenic *E. coli*: ganging up on the host cell. *Curr. Opin. Microbiol.* 12(1), 101–109 (2009).
7. Gomes TA, Elias WP, Scaletsky IC *et al*. Diarrheagenic *Escherichia coli*. *Braz. J. Microbiol.* 47(Suppl. 1), 3–30 (2016).
8. Nougayrede JP, Foster GH, Donnenberg MS. Enteropathogenic *Escherichia coli* effector EspF interacts with host protein Abcf2. *Cell. Microbiol.* 9(3), 680–693 (2007).
9. Weffen AW, Alto NM, Viswanathan VK, Hecht G. *E. coli* secreted protein F promotes EPEC invasion of intestinal epithelial cells via an SNX9-dependent mechanism. *Cell. Microbiol.* 12(7), 919–929 (2010).

10. Aitio O, Hellman M, Skehan B *et al.* Enterohaemorrhagic *Escherichia coli* exploits a tryptophan switch to hijack host f-actin assembly. *Structure* 20(10), 1692–1703 (2012).
11. Holmes A, Muhlen S, Roe AJ, Dean P. The EspF effector, a bacterial pathogen's swiss army knife. *Infect. Immun.* 78(11), 4445–4453 (2010).
12. Kenny B, Valdivia R. Host–microbe interactions: bacteria. *Curr. Opin. Microbiol.* 12(1), 1–3 (2009).
13. Lodato PB, Kaper JB. Post-transcriptional processing of the LEE4 operon in enterohaemorrhagic *Escherichia coli*. *Mol. Microbiol.* 71(2), 273–290 (2009).
14. Berdichevsky T, Friedberg D, Nadler C, Rokney A, Oppenheim A, Rosenshine I. Ler is a negative autoregulator of the LEE1 operon in enteropathogenic *Escherichia coli*. *J. Bacteriol.* 187(1), 349–357 (2005).
15. Nagai T, Abe A, Sasakawa C. Targeting of enteropathogenic *Escherichia coli* EspF to host mitochondria is essential for bacterial pathogenesis: critical role of the 16th leucine residue in EspF. *J. Biol. Chem.* 280(4), 2998–3011 (2005).
16. Nougayrede J, Donnenberg MS. Enteropathogenic *Escherichia coli* EspF is targeted to mitochondria and is required to initiate the mitochondrial death pathway. *Cell. Microbiol.* 6(11), 1097–1111 (2004).
17. Marchès O, Batchelor M, Shaw RK *et al.* EspF of enteropathogenic *Escherichia coli* binds sorting nexin 9. *J. Bacteriol.* 188(8), 3110–3115 (2005).
18. Nieto-Pelegrin E, Kenny B, Martinez-Quiles N. Nck adaptors, besides promoting N-WASP mediated actin-nucleation activity at pedestals, influence the cellular levels of enteropathogenic *Escherichia coli* Tir effector. *Cell. Adhes. Migr.* 8(4), 404–417 (2014).
19. Viswanathan VK, Lukic S, Koutsouris A, Miao R, Muza MM, Hecht G. Cytokeratin 18 interacts with the enteropathogenic *Escherichia coli* secreted protein F (EspF) and is redistributed after infection. *Cell. Microbiol.* 6(10), 987–997 (2004).
20. Shaw RK, Cleary J, Murphy MS, Frankel G, Knutton S. Interaction of enteropathogenic *Escherichia coli* with human intestinal mucosa: role of effector proteins in brush border remodeling and formation of attaching and effacing lesions. *Infect. Immun.* 73(2), 1243–1251 (2005).
21. Muza-Moons MM, Schneeberger EE, Hecht GA. Enteropathogenic *Escherichia coli* infection leads to appearance of aberrant tight junction strands in the lateral membrane of intestinal epithelial cells. *Cell. Microbiol.* 6(8), 783–793 (2004).
22. Quitard S, Dean P, Maresca M, Kenny B. The enteropathogenic *Escherichia coli* EspF effector molecule inhibits PI-3 kinase-mediated uptake independently of mitochondrial targeting. *Cell. Microbiol.* 8(6), 972–981 (2006).
23. Hu C, Chinenov Y, Kerppola TK. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol. Cell.* 9(4), 789–798 (2002).
24. Wilson CG, Magliery TJ, Regan L. Detecting protein–protein interactions with GFP-fragment reassembly. *Nat. Methods* 1(3), 255–262 (2004).
25. Lamesch P, Li N, Milstein S *et al.* hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes. *Genomics* 89(3), 307–315 (2007).
26. Human ORFeome V7.1. <http://horfdb.dfci.harvard.edu/hv7/index.php?page=orfsearch>
27. Chen LY, Liu D, Songyang Z. Telomere maintenance through spatial control of telomeric proteins. *Mol. Cell Biol* 27(16), 5898–5909 (2007).
28. Erlich Y, Mitra PP, DelaBastide M, McCombie WR, Hannon GJ. Alta-cyclic: a self-optimizing base caller for next-generation sequencing. *Nat. Methods* 5(8), 679–682 (2008).
29. Hansen KD, Brenner SE, Dudoit S. Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Res.* 38(12), e131 (2010).
30. DAVID Bioinformatics Resources 6.8. <http://david.ncifcrf.gov/home.jsp>
31. STRING: functional protein association networks. <http://string-db.org/>
32. Jiang L, Schlesinger F, Davis CA *et al.* Synthetic spike-in standards for RNA-seq experiments. *Genome Res.* 21(9), 1543–1551 (2011).
33. Cock PJA, Fields CJ, Goto N, Heuer ML, Rice PM. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res.* 38(6), 1767–1771 (2010).
34. Yan L, Yang M, Guo H *et al.* Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* 20(9), 1131–1139 (2013).
35. Kanopka A. Cell survival: interplay between hypoxia and pre-mRNA splicing. *Exp. Cell. Res.* 356(2), 187–191 (2017).
36. Dean P, Scott JA, Knox AA, Quitard S, Watkins NJ, Kenny B. The enteropathogenic *E. coli* effector EspF targets and disrupts the nucleolus by a process regulated by mitochondrial dysfunction. *PLoS Pathog.* 6(6), e1000961 (2010).
37. Sanchez-Villamil J, Navarro-Garcia F. Role of virulence factors on host inflammatory response induced by diarrheagenic *Escherichia coli* pathotypes. *Future Microbiol.* 10(6), 1009–1033 (2015).
38. McNamara BP, Koutsouris A, O'Connell CB, Nougayrede JP, Donnenberg MS, Hecht G. Translocated EspF protein from enteropathogenic *Escherichia coli* disrupts host intestinal barrier function. *J. Clin. Invest.* 107(5), 621–629 (2001).

39. Alto NM, Weflen AW, Rardin MJ *et al*. The type III effector EspF coordinates membrane trafficking by the spatiotemporal activation of two eukaryotic signaling pathways. *J. Cell Biol.* 178(7), 1265–1278 (2007).
40. Guttman JA, Li Y, Wickham ME, Deng W, Vogl AW, Finlay BB. Attaching and effacing pathogen-induced tight junction disruption *in vivo*. *Cell. Microbiol.* 8(4), 634–645 (2006).
41. Hayward RD, Leong JM, Koronakis V, Campellone KG. Exploiting pathogenic *Escherichia coli* to model transmembrane receptor signalling. *Nat. Rev. Microbiol.* 4(5), 358–370 (2006).
42. Dean P, Kenny B. Intestinal barrier dysfunction by enteropathogenic *Escherichia coli* is mediated by two effector molecules and a bacterial surface protein. *Mol. Microbiol.* 54(3), 665–675 (2004).
43. Dean P, Mühlen S, Quitard S, Kenny B. The bacterial effectors EspG and EspG2 induce a destructive calpain activity that is kept in check by the co-delivered Tir effector. *Cell. Microbiol.* 12(9), 1308–1321 (2010).
44. Thanabalasuriar A, Koutsouris A, Weflen A, Mimeo M, Hecht G, Gruenheid S. The bacterial virulence factor NleA is required for the disruption of intestinal tight junctions by enteropathogenic *Escherichia coli*. *Cell. Microbiol.* 12(1), 31–41 (2010).
45. Peralta-Ramirez J, Hernandez JM, Manning-Cela R *et al*. EspF interacts with nucleation-promoting factors to recruit junctional proteins into pedestals for pedestal maturation and disruption of paracellular permeability. *Infect. Immun.* 76(9), 3854–3868 (2008).
46. Graifer D, Malygin A, Zharkov DO, Karpova G. Eukaryotic ribosomal protein S3: a constituent of translational machinery and an extraribosomal player in various cellular processes. *Biochimie* 99, 8–18 (2014).
47. Warner JR, McIntosh KB. How common are extra-ribosomal functions of ribosomal proteins? *Mol. Cell* 34(1), 3–11 (2009).
48. Tan C, Du X. KRAS mutation testing in metastatic colorectal cancer. *World J. Gastroenterol.* 18(37), 5171–5180 (2012).
49. Yu S, Wang T, Au L. Specific repression of mutant K-RAS by 10–23 DNzyme: sensitizing cancer cell to anti-cancer therapies. *Biochem. Biophys. Res. Commun.* 378(2), 230–234 (2009).
50. Chang F, Steelman LS, Lee JT *et al*. Signal transduction mediated by the Ras/Raf/MEK/ERK pathway from cytokine receptors to transcription factors: potential targeting for therapeutic intervention. *Leukemia* 17(7), 1263–1293 (2003).
51. Friday BB, Adjei AA. K-ras as a target for cancer therapy. *Biochim. Biophys. Acta* 1756(2), 127–144 (2005).
52. Forrester K, Almoguera C, Han K, Grizzle WE, Perucho M. Detection of high incidence of K-ras oncogenes during human colon tumorigenesis. *Nature* 327(6120), 298–303 (1987).
53. Smakman N, Veenendaal LM, van Diest P *et al*. Dual effect of Kras(D12) knockdown on tumorigenesis: increased immune-mediated tumor clearance and abrogation of tumor malignancy. *Oncogene* 24(56), 8338–8342 (2005).
54. Bonnet M, Buc E, Sauvanet P *et al*. Colonization of the human gut by *E. coli* and colorectal cancer risk. *Clin. Cancer Res.* 20(4), 859–867 (2014).
55. Khan S. Potential role of *Escherichia coli* DNA mismatch repair proteins in colon cancer. *Crit. Rev. Oncol. Hematol.* 96(3), 475–482 (2015).
56. Mármol I, Sánchez-de-Diego C, Pradilla Dieste A, Cerrada E, Rodríguez Yoldi M. Colorectal carcinoma: a general overview and future perspectives in colorectal cancer. *Int. J. Mol. Sci.* 18(1), 197 (2017).
57. Maddocks OD, Scanlon KM, Donnenberg MS. An *Escherichia coli* effector protein promotes host mutation via depletion of DNA mismatch repair proteins. *MBio* 4(3), e113–e152 (2013).
58. Dean P, Kenny B. A bacterial encoded protein induces extreme multinucleation and cell–cell internalization in intestinal cells. *Tissue Barriers* 1(1), e22639 (2013).
59. Tahoun A, Siszler G, Spears K *et al*. Comparative analysis of EspF variants in inhibition of *Escherichia coli* phagocytosis by macrophages and inhibition of *E. coli* translocation through human- and bovine-derived M cells. *Infect. Immun.* 79(11), 4716–4729 (2011).
60. Grewal T, Hoque M, Conway J *et al*. Annexin A6-A multifunctional scaffold in cell motility. *Cell Adh. Migr.* 11(3), 288–304 (2017).
61. Koese M, Rentero C, Grewal T. Annexin A6 is a scaffold for PKCa to promote EGFR inactivation. *Oncogene* 32(23), 2858–2872 (2013).
62. Roxas JL, Ryan K, Vedantam G, Viswanathan VK. Enteropathogenic *Escherichia coli* dynamically regulates EGFR signaling in intestinal epithelial cells. *Am. J. Physiol. Gastrointest. Liver Physiol.* 307(3), G374–G380 (2014).
63. Clements A, Stoneham CA, Furniss RCD, Frankel G. Enterohaemorrhagic *Escherichia coli* inhibits recycling endosome function and trafficking of surface receptors. *Cell. Microbiol.* 16(11), 1693–1705 (2014).
64. Canu N, Pagano I, La Rosa LR *et al*. Association of TrkA and APP is promoted by NGF and reduced by cell death-promoting agents. *Front. Mol. Neurosci.* 10, 15 (2017).
65. Yue L, Li L, Li D *et al*. High-throughput screening for survivin and borealin interaction inhibitors in hepatocellular carcinoma. *Biochem. Biophys. Res.* 484(3), 642–647 (2017).
66. Wu W, Hsu YS, Bi L, Songyang Z, Lin X. CARD9 facilitates microbe-elicited production of reactive oxygen species by regulating the LyGDI-Rac1 complex. *Nat. Immunol.* 10(11), 1208–1214 (2009).
67. Huang J, Craggs TD, Christodoulou J, Jackson SE. Stable intermediate states and high energy barriers in the unfolding of GFP. *J. Mol. Biol.* 370(2), 356–371 (2007).

68. Lee OH, Kim H, He Q *et al.* Genome-wide YFP fluorescence complementation screen identifies new regulators for telomere signaling in human cells. *Mol. Cell. Proteomics* 10(2), M110–M1628 (2011).
69. Viswanathan VK, Koutsouris A, Lukic S *et al.* Comparative analysis of EspF from enteropathogenic and enterohemorrhagic *Escherichia coli* in alteration of epithelial barrier function. *Infect. Immun.* 72(6), 3218–3227 (2004).
70. Miranda M, Escote X, Ceperuelo-Mallafre V *et al.* Paired subcutaneous and visceral adipose tissue aquaporin-7 expression in human obesity and Type 2 diabetes: differences and similarities between depots. *J. Clin. Endocrinol. Metab.* 95(7), 3470–3479 (2010).
71. Kozono D, Yasui M, King LS, Agre P. Aquaporin water channels: atomic structure molecular dynamics meet clinical medicine. *J. Clin. Invest.* 109(11), 1395–1399 (2002).
72. Witke W. The role of profilin complexes in cell motility and other cellular processes. *Trends Cell Biol.* 14(8), 461–469 (2004).
73. Llado A, Timpson P, Vila DMS *et al.* Protein kinase C δ and calmodulin regulate epidermal growth factor receptor recycling from early endosomes through Arp2/3 complex and cortactin. *Mol. Biol. Cell* 19(1), 17–29 (2008).
74. Smythe E, Ayscough KR. Actin regulation in endocytosis. *J. Cell. Sci.* 119(22), 4589–4598 (2006).
75. Da Costa SR, Okamoto CT, Hamm-Alvarez SF. Actin microfilaments *et al.* – The many components, effectors and regulators of epithelial cell endocytosis. *Adv. Drug Deliv. Rev.* 55(11), 1359–1383 (2003).
76. Grewal T, Enrich C. Annexins-modulators of EGF receptor signalling and trafficking. *Cell Signal.* 21(6), 847–858 (2009).
77. Crane JK, Oh JS. Activation of host cell protein kinase C by enteropathogenic *Escherichia coli*. *Infect. Immun.* 65(8), 3277–3285 (1997).
78. Shen-Tu G, Kim H, Liu M, Johnson-Henry KC, Sherman PM. Protein kinase C mediates enterohemorrhagic *Escherichia coli* O157:H7-induced attaching and effacing lesions. *Infect. Immun.* 82(4), 1648–1656 (2014).
79. Celli J, Olivier M, Finlay BB. Enteropathogenic *Escherichia coli* mediates antiphagocytosis through the inhibition of PI 3-kinase-dependent pathways. *EMBO J.* 20(6), 1245–1258 (2001).
80. Roxas JL, Koutsouris A, Viswanathan VK. Enteropathogenic *Escherichia coli*-induced epidermal growth factor receptor activation contributes to physiological alterations in intestinal epithelial cells. *Infect. Immun.* 75(5), 2316–2324 (2007).
81. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 61(5), 759–767 (1990).
82. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. *Cell* 87(2), 159–170 (1996).
83. Smith G, Carey FA, Beattie J *et al.* Mutations in APC, Kirsten-ras, and p53 – alternative genetic pathways to colorectal cancer. *Proc. Natl Acad. Sci. USA* 99(14), 9433–9438 (2002).
84. Choi HJ, Kim J, Do KH, Park S, Moon Y. Enteropathogenic *Escherichia coli*-induced macrophage inhibitory cytokine 1 mediates cancer cell survival: an *in vitro* implication of infection-linked tumor dissemination. *Oncogene* 32(41), 4960–4969 (2013).
85. Magdy A, Elhadidy M, Abd Ellatif ME *et al.* Enteropathogenic *Escherichia coli* (EPEC): does it have a role in colorectal tumorigenesis? A prospective cohort study. *Int. J. Surg.* 18, 169–173 (2015).
86. Maddocks OD, Short AJ, Donnenberg MS, Bader S, Harrison DJ. Attaching and effacing *Escherichia coli* downregulate DNA mismatch repair protein *in vitro* and are associated with colorectal adenocarcinomas in humans. *PLoS ONE* 4(5), e5517 (2009).
87. Vogelmann R, Amieva MR. The role of bacterial pathogens in cancer. *Curr. Opin. Microbiol.* 10(1), 76–81 (2007).
88. Kipanyula MJ, Seke Etet PF, Vecchio L, Farahna M, Nukenine EN, Nwabo Kamdje AH. Signaling pathways bridging microbial-triggered inflammation and cancer. *Cell Signal.* 25(2), 403–416 (2013).