# Identification of a novel interaction between SMC1 DNA damage repair protein and *Escherichia coli* O157: H7 EspF using co-immunoprecipitation combined with mass spectrometry

Muqing Fu[1], Ying Hua[1], Jiali Wu[1], Zhikai Zhang[1], Jinyue Liu[1], Xiaoxia Li[1], Yuting Fang[1], Bao Zhang[1], Wei Zhao[1], Chengsong Wan [1,2*]

[1]Department of Microbiology, School of Public Health, Southern Medical University, Guangzhou, China.

[2] Key Laboratory of Tropical Disease Research of Guangdong Province, Southern Medical University, Guangzhou, China.

* Correspondence: gzwcs@smu.edu.cn. Fax: +86-20-61648324

**Abstract:** The enterohemorrhagic *Escherichia coli* (EHEC) O157: H7 EspF is known to be a multifunctional effector that triggers several damage processes in the host cells. However, in the process of EHEC O157: H7 infection, the interactions between EspF, its N- or C-terminus, and host proteins are still unclear. In this work, we use co-immunoprecipitation combined with mass spectrometry (CoIP-MS) to screen the interactions between EspF/EspF-N/EspF-C terminus and host proteins. A total of 311 host proteins are analyzed. The N-terminus of EspF is found to interact with 192 proteins, whereas 205 proteins interact with the C-terminus. These proteins are mainly involved in RNA splicing, endoplasmic reticulum stress, and a variety of metabolic signaling pathways. Surprisingly, MS results reveal that EspF can also phosphorylate H2AX, suggesting that EspF may directly mediate DNA damage. Here, by western blot and immunofluorescence (IF), we verified that EspF can cause phosphorylation of H2AX and mediates cell multi-nuclearation and cell hypertrophy. Furthermore, we verify here for the first time that SMC1 interacts with EspF -C-terminus, and provide evidence that EspF increases p-SMC1 levels. p-SMC1 is known to influence S-phase cell cycle arrest and usually increases during periods of DNA damage. Our work revealed a novel interaction between EspF and the host protein SMC1 and lays a foundation for further research on EspF-mediated host DNA damage, apoptosis, and even colorectal carcinogenesis.

**Keywords:** Enterohemorrhagic *Escherichia coli* O157: H7, EspF, SMC1, CoIP-MS, DNA damage

## 1. Introduction

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

Enterohemorrhagic *E. coli* (EHEC) O157: H7 is an important foodborne pathogen that causes human diarrhea, hemorrhagic colitis and hemolytic uremic syndrome[1,2](Al, 2005; Kaper, 2005). EHEC O157: H7 adheres to the brush border of epithelial cells and uses T3SS type III secretion system to inject effector proteins into host cells[3]. EspF is one of the most important virulence factors of A/E pathogens[4], it destroys the tight junctions of intestinal epithelial cells[5], leading to cytoskeletal rearrangement and disappearance of intestinal epithelial microvilli[6], and induces host cell apoptosis [7].

The N-terminus of EspF (1–73 aa) has a host cell mitochondrial targeting signal (MTS, 1–24 aa) [8], and a nucleus-binding domain (NTD, 21–74aa)[9]. Roxas found that EspF localizes to mitochondria, destroys mitochondrial membrane potential, and activates the apoptotic proteases 3 and 9. The apoptotic proteases can cleave the epidermal growth factor receptor (EGFR) of host cells, leading to EGFR degradation and a dramatic increase in host cell death in the late stages of infection[10]. The C-terminus (73–248 aa) is composed of four highly homologous proline-rich sequences (PRRs), each containing a protein binding site Src homology 3 (SH3) motif and a neuronal Wiskott–Aldrich syndrome protein (N-WASP) binding domain[11,12].

Although there has been some research on the interaction between EspF and host proteins, the molecular mechanisms of EspF or N-C-terminus interacted with host proteins and the impact of these interactions are still unclear. Identifying the interaction between EspF and the host is a necessary step in elucidating the pathogenic mechanism.

CoIP-MS is one of the most widely used high-throughput techniques for discovering protein–protein interactions (PPI). In our research, we analyzed the molecular function (MF), biological processe (BP), and cellular component (CC) of the related proteins using the Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Clusters of Orthologous Groups (COG) functional annotations. The STRING online tool was used to analyze interactions between target proteins, and Cytoscape software was used to draw the protein–protein interaction network. The interactions were validated by immunoprecipitation, and co-localization was performed by confocal microscopy. Our research provided a protein network map between EspF and the host, laying a foundation for further research on how EspF directly mediates DNA damage in host cells and even causes colorectal carcinogenesis.

## 2. Results

### *2.1. Isolation of differential bands in EspF/EspF-N/EspF-C groups from cell lysates*

After transfecting pEGFP-EspF, pEGFP-EspF/N, and pEGFP-EspF/C-terminus encoding plasmids into 293T cells for 48 hours, the lysis proteins were added to Flag columns and IgG columns for co-immunoprecipitation (Figure 1). pEGFP-EspF was about 60 kDa, pEGFP-EspF/N was 35 kDa, and pEGFP-EspF/C was 55 kDa, as shown in Figure 1A. Compared with the IgG group, between 35–40 kDa, there were two differentially expressed bands in the pEGFP-EspF group. At 40 kDa, there was a differential band in the pEGFP-EspF/N group. At 120 kDa, there was a differential band in the pEGFP-EspF/C group. Besides these bands, the bands at the same position in the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

lanes of the respective IgG groups were also cut out. The bands were digested by mass spectrometry to detect the interacting proteins, and the Flag group-specific proteins (minus the IgG group proteins) were considered to be the putative interacting proteins(Figure 1B).

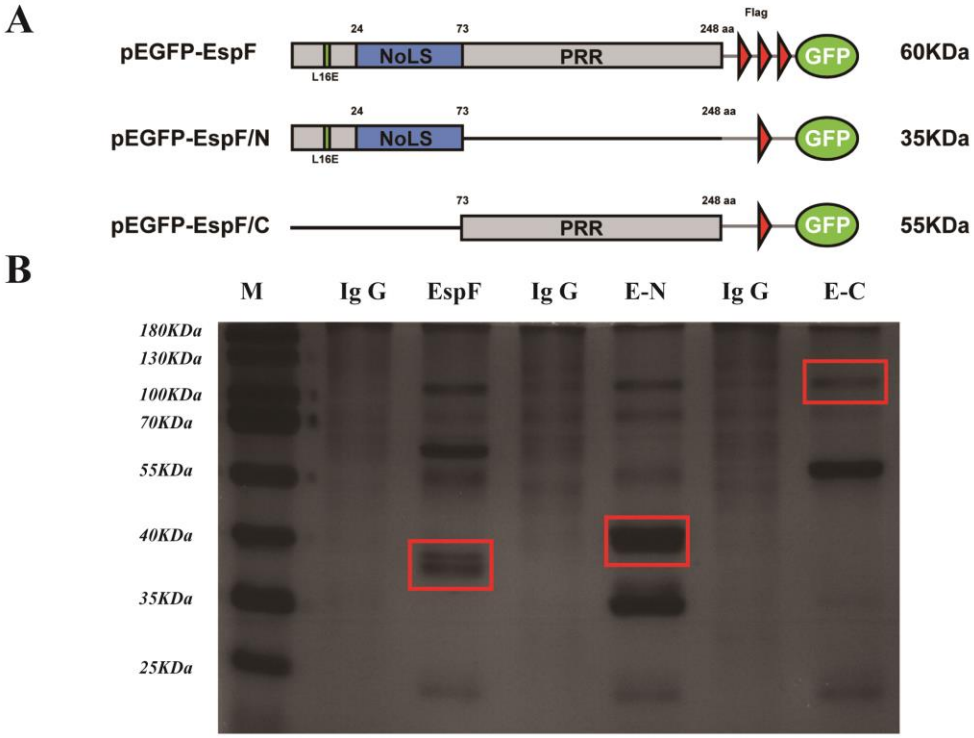



**Figure 1.** Differential bands of pEGFP-EspF/pEGFP-EspF-N(E-N)/pEGFP-EspF-C(E-C) group. (A) Diagram of EspF /EspF-N/EspF-C plasmid constructs used in this study. pEGFP-EspF, pEGFP-EspF/N, and pEGFP-EspF/C are tagged with Flag. aa: Amino acid. (B) Silver staining of the gel. pEGFP-EspF, pEGFP-EspF/N, and pEGFP-EspF/C were transfected into 293T cells. From left to right is Maker, pEGFP-EspF IgG group and Flag group, pEGFP-EspF/N IgG group and Flag group, and pEGFP-EspF/C IgG group and Flag group are shown. The proteins were lysed and electroporated on 10% SDS-PAGE. The differential bands are shown in the figure with red textboxs.

### *2.2. Prediction and analysis of protein interactions with EspF*

A total of 708 proteins were identified in this work, of which 311, 192, and 205 proteins were detected in the pEGFP-EspF group, pEGFP-EspF/N group, and pEGFP-EspF/C group respectively. All possible target proteins that interacted with EspF were loaded into the DAVID database for KEGG pathway annotation and GO enrichment. The threshold was set to ps0.05, and pathways or gene functions with higher counts were analyzed. The top 15 proteins were plotted with Graphpad Prism 6 (Table 1). Analysis of the differential bands at about 38 kDa in the pEGFP-EspF group by GO annotation analysis revealed that the interacting proteins were involved in 25 BPs. Of these, the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

primary BPs were intracellular processes (12.3%), and metabolic processes (10.9%). CCs mainly involved cellular anatomical entities (37.7%), intracellular (36.6%), and protein-containing complex (20.9%). MFs mostly involved binding (53.6%) and catalytic activity (22.3%) (Figure 2A). Pathway analysis showed that the interacting proteins were notably involved in metabolic pathways (32.6%), carbon metabolism pathways (19.6%), and biosynthesis of amino acids (13.0%) (Figure 2B).

We used STRING to analyze the target proteins with which EspF interacted, and found that RPS6, RPL14, and EIF2S1 had the highest connectivity (Figure 2C). RPS6 plays an essential role in controlling cell growth and proliferation by selectively translating specific kinds of mRNAs [13]. RPL14 is a large ribosomal subunit component that plays a role in mRNA catabolism and translation [14]. EIF2S1 works in the early stages of protein synthesis by forming a ternary complex with GTP and initiator tRNA [15]. This analysis showed that in addition to gene transcription regulation and protein synthesis, EspF also plays a crucial role in cell proliferation and catabolism. The GO annotation analysis of the 36 kDa differential band showed CCs and MFs results were consistent with the results of the 38 kDa differential band (Figure S1A).

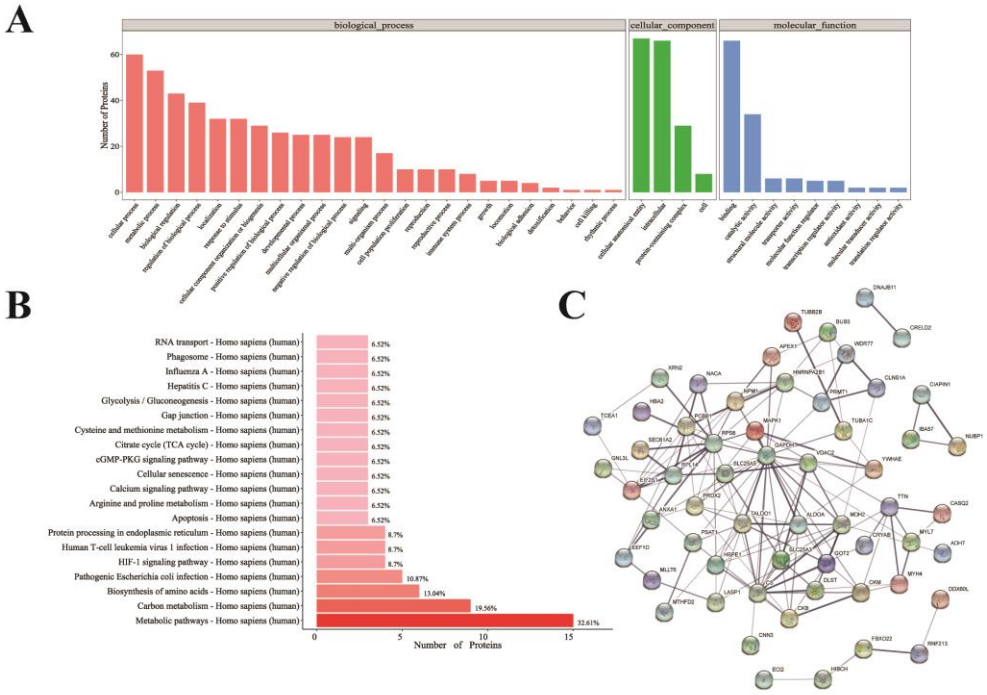



**Figure 2.** GO enrichment, KEGG pathway, and STRING analyses of the 38 kDa proteins interacting with EspF. (A) GO annotation of EspF-interacting proteins in terms of biological processes, cellular components, and molecular functions. (B) The KEGG analysis of the distribution of the top 20 protein interaction pathways. (C) STRING analysis of statistically significant proteins associated with EspF interaction. The networks also illustrate the functional relationships (the edges) between the nodes; the thickness of the nodes is directly proportional to the association's significance score.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

*2.3. Prediction and analysis of proteins that interacted with EspF-N or C-terminus*

The GO annotation analysis of the differential bands at about 40 kDa in the pEGFP-EspF/N group showed that 192 interacting proteins were involved in 25 BPs, such as intracellular processes (12.3%), and metabolic processes (10.8%). Pathway analysis revealed that the interacting proteins were mainly involved in metabolic pathways (27.1%), protein processing in the endoplasmic reticulum (9.41%), and pathways in cancer (8.2%) (Figure 3AB). STRING analysis of target proteins interacting with the EspF-N terminus also found that RPL8, RPS9, and EIF3I proteins had the highest degree of connectivity, indicating that EspF may use its N-terminus to recognize ribosomes binding sites (Figure 3C).

GO annotation analysis of pEGFP-EspF/C group in the 120 kDa differential band showed that BPs enrichment mostly involved translation (23.1%), and the oxidation-reduction process (15.4%). CCs mainly involved intermediate filaments (21.3%) and ribosomes (13.5%). MFs mostly involved protein binding (22.4%) and ATP binding (18.6%) (Figure 4A). Of the interacting proteins, 30.3% were localized in the cytoplasm, and 27.27% were localized in the nucleus (Figure 4B). COG analysis showed that interacting proteins mostly involved translation, ribosomal structure, and posttranslational modification (Figure 4C). KEGG analysis showed that most of the cellular processes involved protein transport, as well as signal pathway that affected cell growth and apoptosis (Figure 4D). Among these proteins, TUBB, and ANXA2 showed strong interaction with EspF-C. TUBB, a main component of tubulin, demonstrated GTPase activity, and plays a key role in microtubule cytoskeleton organization [16]. Studies have shown that EspF interacts with SNX9 to induce the formation of membrane tubules and the host cell membrane change [17]. In addition to the SNX9 binding site, EspF can also activate N-WASP to induce actin polymerization and TJ disruption [18]. Microfilaments, microtubules, and intermediate fibers together make up the cytoskeletal system. Therefore, we speculated that EspF might also cause cytoskeleton rearrangement through binding TUBB.

In our previous research, ANXA6 and EspF were confirmed to interact [19]. In this study, we found ANXA2, a member of the ANXA protein family, which is a calcium-regulated membrane-bound protein. ANXA2 participates in the heat stress response by interacting with Hsp90 [20]. Moreover, EspF may regulate calcium ion accumulation and calcium channel protein activity by interacting with ANXA2.

In addition, we also found SMC1. The interaction between EspF and SMC1 was also noted in our previous research, and involves multiple signaling pathways [19]. It is understood that SMC1 is a chromosomal structural protein that plays a role in DNA replication and cohesion of sister chromatids. It involves in chromosome dynamics, cell cycle regulation, cell proliferation, and genome stability [21]. Furthermore, when cells are stimulated by DNA damage, SMC1 can be phosphorylated by ATM or ATR to participate in DNA repair, and acts as a downstream effector in the ATM/NBS1 and ATR/MSH2 branches to activate the S-phase checkpoint [22]. Hence, we hypothesized that EspF could mediate cellular DNA damage repair by phosphorylating SMC1.

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

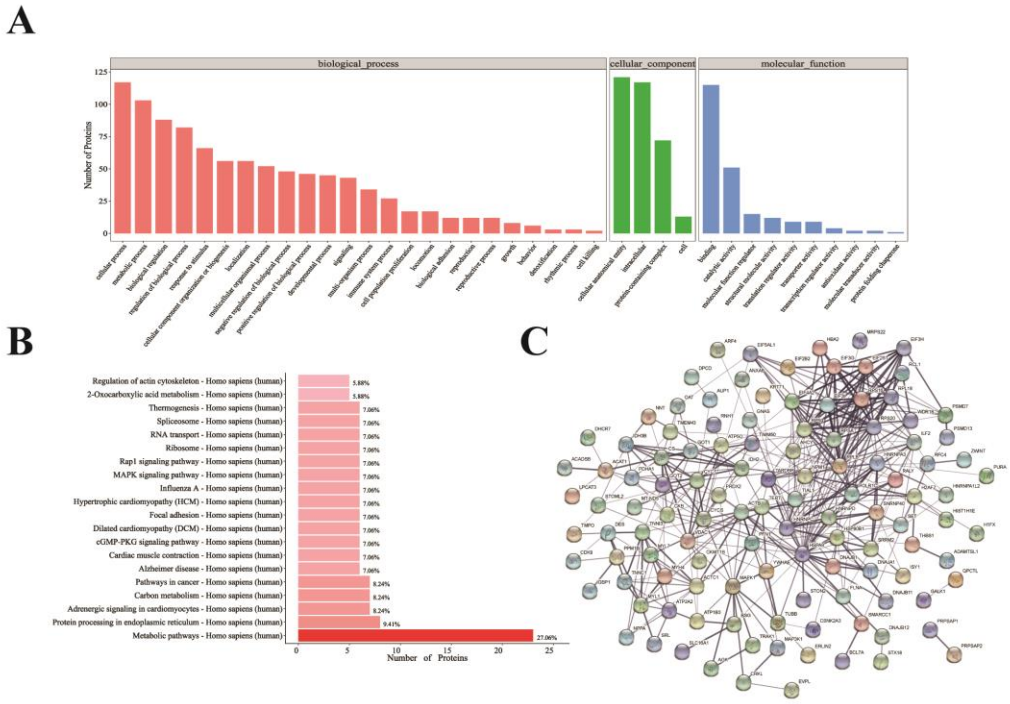



**Figure 3.** GO enrichment, KEGG pathway, and STRING analyses of the 40 kDa proteins that interacted with the EspF-N terminus. (A) GO annotation of identified proteins that interacted with the EspF-N terminus in terms of BPs, CCs, and MFs. (B) KEGG analysis of the distribution of the top 20 protein interaction pathways. (C) STRING analysis of statistically significant proteins detected in proteins interacting with EspF-N.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

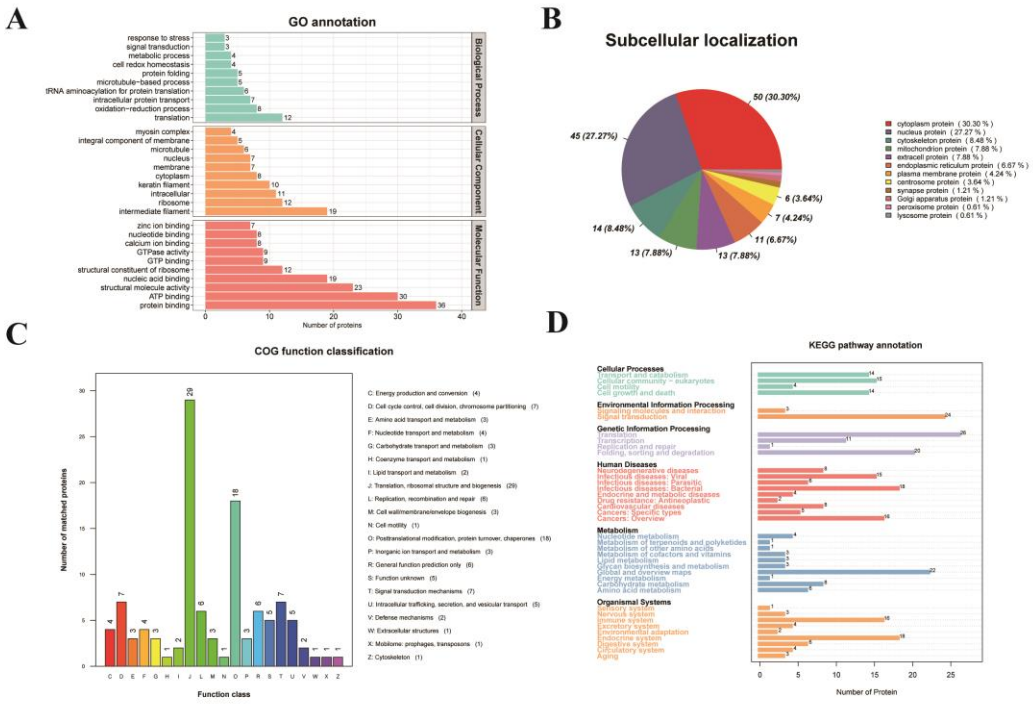



**Figure 4.** GO enrichment, KEGG pathway, and STRING analyses of the 130 kDa proteins that interacted with the EspF-C terminus. (A) GO annotation of the identified EspF-C-interacting proteins in terms of BPs, CCs, and MFs. (B) Subcellular location of the interacting proteins. (C) KEGG analysis of the distribution of the interacting protein pathways. (D) The COG function classification of the interacting protein (COG) pathways.

**Figure 5.** EspF can cause phosphorylation of H2AX and mediates cell multi-nucleation and cell hypertrophy.（A）Caco2 cells treated with strains for 6 h, then p-H2AX protein expression were assessed by western blotting. ACTIN was used as loading control. Relative protein levels were quantified using ImageJ software. Data are shown as the mean±SD of three independent repeats. **P* < 0.05, ****P* < 0.01. (B) Immunofluorescence microscopy of p-H2AX protein. The nucleus was labeled with DAPI in blue. p-H2AX was labeled with anti-p-H2AX antibody in red. Images were acquired on an FV1000 confocal laser scanning microscope using a 63X oil objective. Quantification analysis of fluorescence intensity of p-H2AX by ImageJ software. Error bars represent means±SD from three independent experiments. **P* < 0.05, ****P* < 0.01. (C) Immunofluorescence microscopy of CK18 protein. Caco2 cells transfected with plasmids for 48 h, The nucleus was labeled with DAPI in blue. CK18 was labeled with anti-CK18 antibody in red. Green indicates fluorescence expressed by the fluorescent plasmid. Images were acquired on an FV1000 confocal laser scanning microscope using a 63X oil objective.

*2.4. EspF may mediate DNA damage by modifying histones*

 doi:10.20944/preprints202012.0551.v1

Mass spectrometry results showed that EspF could modify multiple sites of various proteins (Table S2). Among the modification results, HAX1, EIF3I, ATG16L1, and DNA damage-binding proteins had high scores. HAX1 recruits the Arp2/3 complex to the cell cortex, and through its interaction with KCNC3, it reconstitutes the cortical actin cytoskeleton [23]. Previous studies have shown that EspF cooperates with Arp2/3, profifilin, actin, and ZO-1 to redistribute actin and potentially disrupt the TJs.

EspF may rearrange the cytoskeleton through phosphorylation and acetylation of the upstream HAX1. EspF also phosphorylate some autophagy-related proteins, including ATG16L1 and PHB, that mediate such as cell autophagy and protein transport signals[24].

We also discovered that EspF also modified damage-specific DNA-binding protein 2 (DDB2), a type of DNA repair protein. DDB2 is originally identified as a DNA damage-recognition factor that promotes genomic nucleotide excision repair (GG-NER) in human cells. Previous research showed that EspF could cause cell apoptosis by targeting mitochondria and releasing cytochrome c [25], but it is still unclear whether EspF can mediate DNA damage through modification of histones, which lays the foundation for our future research.

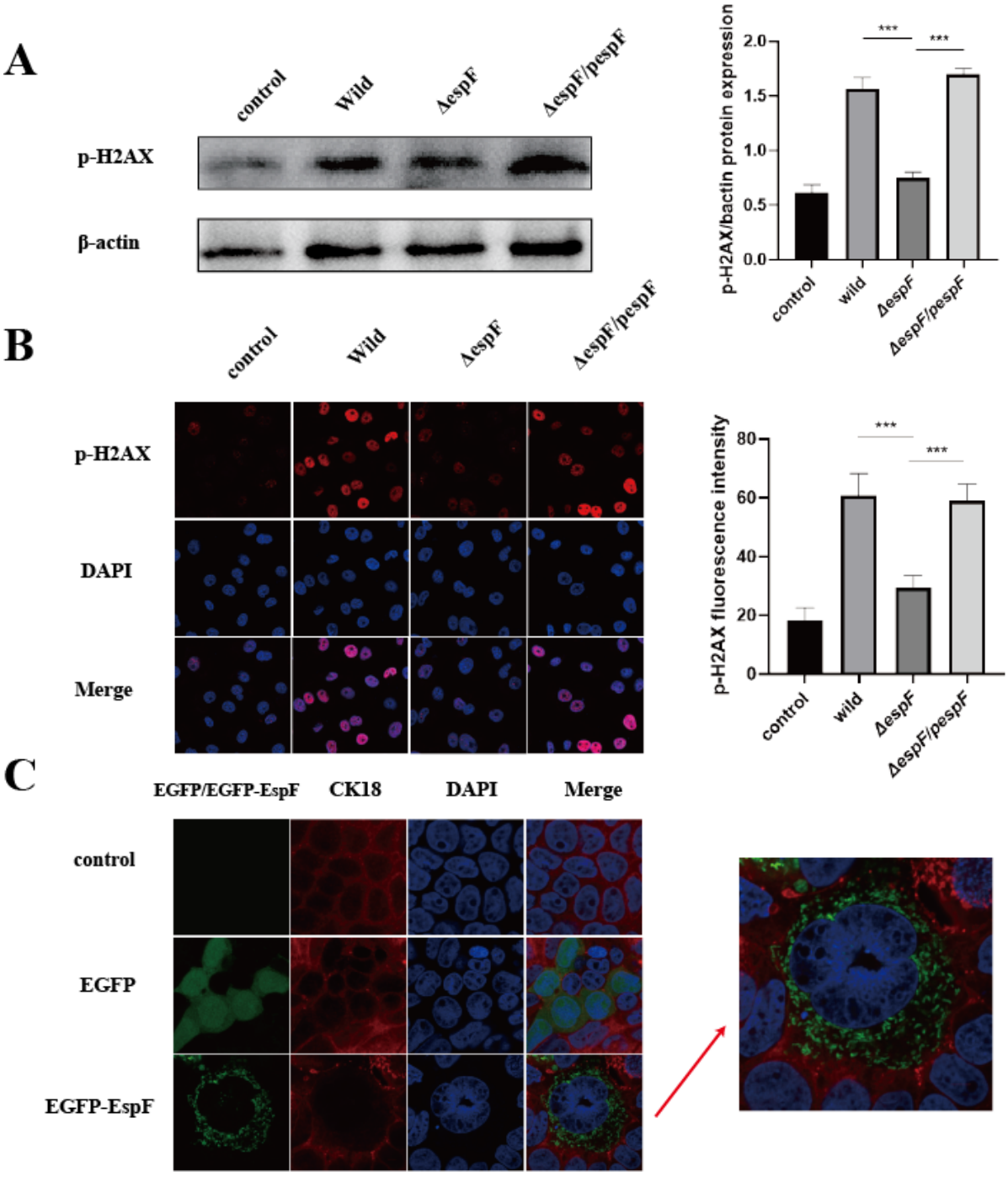
Next, we also discovered that EspF may affect epigenetic changes in host cells, it could modify a series of histones, such as Histone H1.3, Histone H2A type 1, and Histone H1x, which are the core components of nucleosomes. Nucleosomes encapsulate DNA into chromatin, limiting DNA entry into cellular mechanisms [26]. Therefore, histone modification plays a central role in transcription regulation, DNA damage repair, DNA replication, and chromosomal stability. Importantly, we found that EspF could phosphorylate histone H2AX, a known marker of DNA damage [27], suggesting that EspF can cause DNA damage.

To determine whether EspF can directly cause DNA damage and the role in this process, we evaluated the level of p-H2AX in Caco2 cells by infecting with strains. Compared with *ΔespF* group, the levels of p-H2AX in the wild group were significantly increased. Then we infected Caco2 with strains and then observed the distribution of p-H2AX by immunofluorescence. As shown in Figure 5, in non-infected Caco-2 cells (NC), p-H2AX protein was less expression. In the wild group, the fluorescence of p-H2AX appeared strongly positive in the nucleus. Instead, *ΔespF* group weakened fluorescence signals and dotted around the nucleus. The fluorescence of p-H2AX was compensated by the complement of *espF*. The fluorescence intensity of p-H2AX was quantified by ImageJ software. The signal intensity of p-H2AX in the wild groups were tripled than *ΔespF* group.

Morphological changes of cells are also one of the manifestations of DNA damage, EPEC EspF induces extreme multi-nucleation in small intestinal epithelial cells, these extreme phenotypes were dependent on a C-terminal polyproline-rich domain of EspF [28]. In addition, EspF collaborates with cytokeratin 18 and 14-3-3ζ, resulting in the alteration of the intermediate filament network[29]. we hypothesized that EHEC EspF can also cause cells multi-nucleation. In our study, Caco2 cells were transfected with

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

EGFP-EspF for 4 days and stained with anti-CK18 antibody, Cells expressing EGFP-EspF on day 2 post-transfection remained mononuclear. From day 2 to day 4 post-transfection, Cells expressing EspF-GFP exhibited multi-nucleation and hypertrophy (counted as three or more nuclei). Meanwhile, in cells expressing EGFP and NC group, CK18 was distributed continuously along the cell membrane, with a complete structure and clear boundaries, compared with uninfected monolayers, the reticulate IF net-work seen in transfected with EGFP-EspF was collapsed, with discontinuous distribution and weakened fluorescence signals, as shown in Figure 5C. These results indicate that EspF can cause DNA damage and can destroy IF network,which lead to morphological changes, such as cell multi-nucleation and cell hypertrophy.



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

**Figure 5.** EspF can cause phosphorylation of H2AX and mediates cell multi-nucleation and cell hypertrophy.（A）Caco2 cells treated with strains for 6 h, then p-H2AX protein expression were assessed by western blotting. ACTIN was used as loading control. Relative protein levels were quantified using ImageJ software. Data are shown as the mean±SD of three independent repeats. **P* < 0.05, ****P* < 0.01. (B) Immunofluorescence microscopy of p-H2AX protein. The nucleus was labeled with DAPI in blue. p-H2AX was labeled with anti-p-H2AX antibody in red. Images were acquired on an FV1000 confocal laser scanning microscope using a 63X oil objective. Quantification analysis of fluorescence intensity of p-H2AX by ImageJ software. Error bars represent means±SD from three independent experiments. **P* < 0.05, ****P* < 0.01. (C) Immunofluorescence microscopy of CK18 protein. Caco2 cells transfected with plasmids for 48 h, The nucleus was labeled with DAPI in blue. CK18 was labeled with anti-CK18 antibody in red. Green indicates fluorescence expressed by the fluorescent plasmid. Images were acquired on an FV1000 confocal laser scanning microscope using a 63X oil objective.

### *2.5. SMC1 was identified as a novel EspF-interacting protein, and its interaction via EspF-C terminus*

In our previous research, we identified SMC1 that binds to EspF, and involves in DNA damage repair[19]. MS results also revealed EspF-C interacted with SMC1 protein. To our knowledge, we are the first to study the mechanism of interaction between EspF and DNA damage repair proteins SMC1. Confocal analysis and Co-IP further confirmed the interaction between EspF and SMC1. We transfected EGFP/EGFP-EspF into Caco2 cells, immunofluorescence results showed that SMC1 was mostly distributed in the nucleus in control group, while after transfecting EspF, SMC1 and EspF were apparently co-localized in the cytoplasm. Then we infected Caco2 cells with strains, compared with cells infected by *ΔespF*, SMC1 was distributed more from the nucleus to the cytoplasm in the cells infected with EHEC, suggesting that SMC1 may not play its usual role in the nucleolus. Also, the phenotype of SMC1 of cells infected with *ΔespF/pespF* was consistent with that infected with EHEC (Figure 6AB).

Co-immunoprecipitation results confirmed that EspF interacted with SMC1. We transfected EGFP-EspF/EGFP-EspF-N/EGFP-EspF-C into 293T cells, and then obtained the cell lysate. As shown in Figure 7A, we observed EspF/EspF-C-terminal protein co-precipitated with SMC1, whereas the N-terminal protein interacted very weakly with SMC1. Thus, we confirmed EspF interacted with SMC1 by its C-terminus.

To investigate the interaction between EspF and SMC1, we measured SMC1 and its phosphorylation levels after pEGFP-EspF transfected into Caco2 cells. The level of SMC1 remained unchanged after transfection with EspF, but p-SMC1 significantly increased (Figure 7B). We then verified the results by infecting HT-29 cells with the strain. The expression of p-SMC1 in cells transfected with EHEC was higher than in cells transfected with the *ΔespF* strain, and infecting the complement *ΔespF/pespF*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

restored p-SMC1 expression(Figure 7C). The above results verified that EspF could increase the expression level of p-SMC1.

These findings suggest that EspF interacts the host SMC1 protein through its C-terminus, we speculated that EspF lead to DNA damage, which stimulates the S-phase checkpoint by increasing p-SMC1 expression, thus mediating damage repair. The role of EspF-SMC1 in EHEC pathogenicity remains to be elucidated.

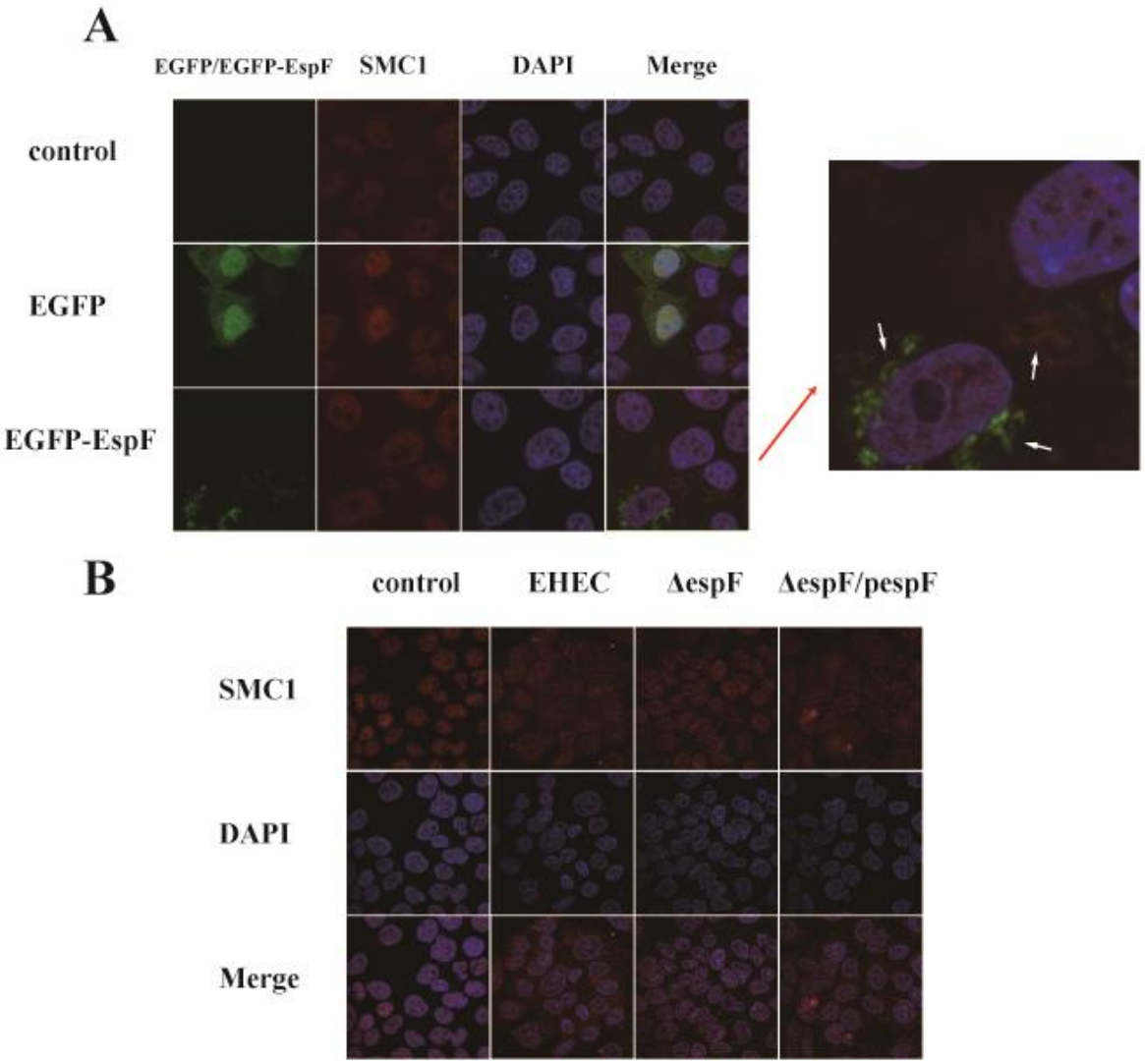



**Figure 6.** EspF and SMC1 transiently colocalized, and EspF relocalizes SMC1 into the cytoplasm. A. Caco2 cells were grown at 70% of confluence in mock conditions and transfected with EGFP-N1/EGFP-EspF. After 48 h, cells were fixed, permeabilized, and immunostained with anti-SMC1 antibody followed by a secondary antibody (red), Slides were analyzed and recorded by confocal microscopy (63X zoom 3). The arrow points to co-location. B. Caco2 cells were grown at 70% of confluence in mock conditions and infected with EHEC/EHEC*ΔespF* /EHEC*ΔespF*+ *espF* at a MOI of 100:1 for 6 h. Then, cells underwent immunofluorescence as shown above.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

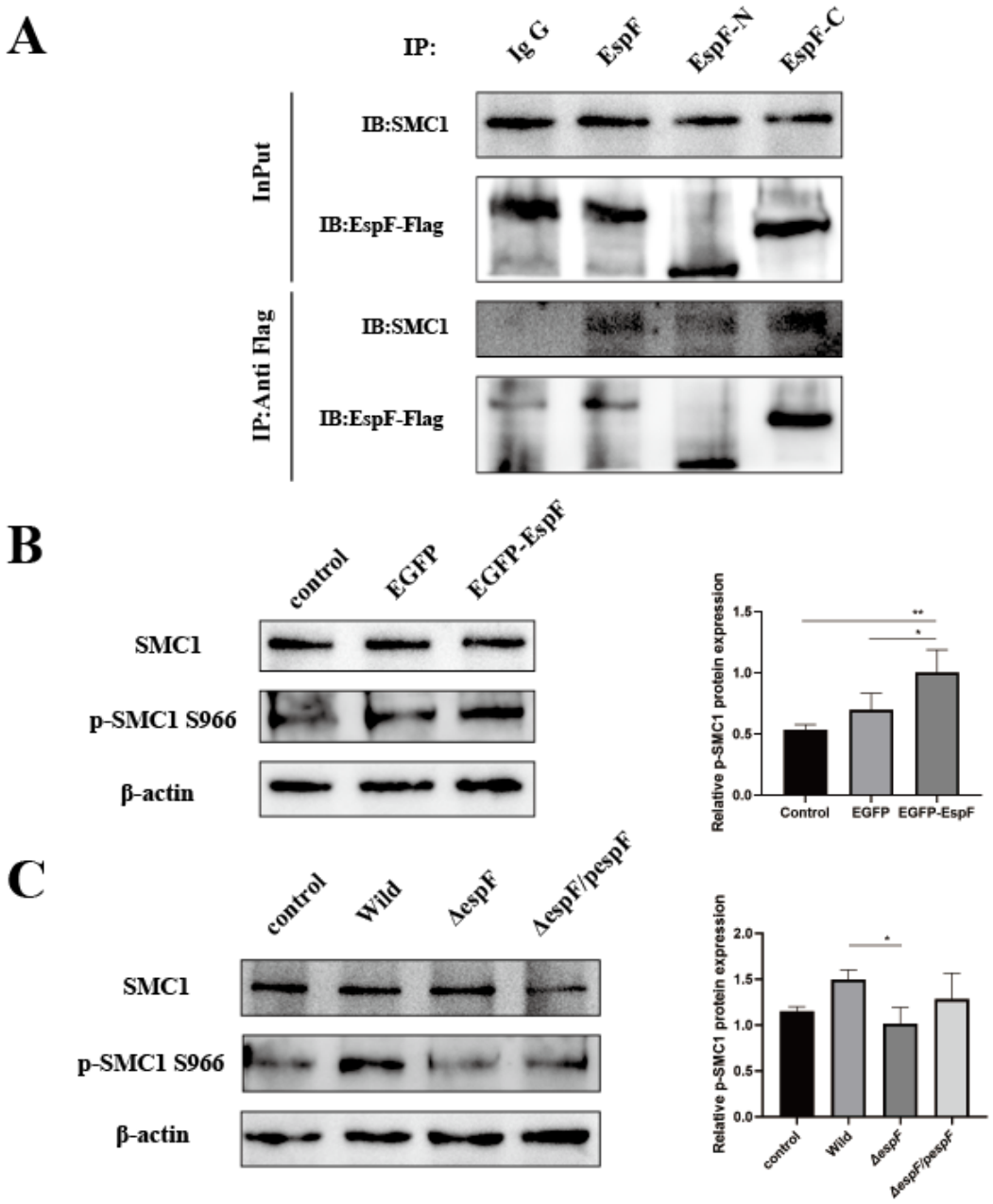



**Figure 7.** EspF interacts with SMC1 and can increase its phosphorylation level. （A）Co-immunoprecipitation was performed to identify the interactions between SMC1 and EspF/EspF-N/EspF-C. pEGFP-EspF, pEGFP-EspF-N, and pEGFP-EspF-C were transfected into 293T cells. Cell proteins were extracted and incubated with Flag/IgG agarose beads. Then, the cell proteins (input) and the bead binding proteins were prepared for 10% SDS-PAGE. (B) Caco2 were transfected with pEGFP-N1/pEGFP-EspF for 48 h, after which SMC1 and p-SMC1 expression were assessed by western blotting. ACTIN was used as loading control. Relative protein levels were quantified using ImageJ software. Data are shown as the mean±SD of three independent repeats.(C) HT-29 treated with strains for 6 h, after which SMC1 and p-

 doi:10.20944/preprints202012.0551.v1

SMC1 expression were assessed by western blotting. ACTIN was used as loading control. Relative protein levels were quantified using ImageJ software. Data are shown as the mean±SD of three independent repeats.

## 3. Discussion

Previous studies screened interactions between EspF and host proteins through yeast two-hybrid and BiFC fluorescence methods [19,30]. The former cannot determine whether the interactions exist in mammalian cells, while the latter uses tags that may affect the conformation or activity of the mark protein, leading to false-negative results. Therefore, we opted for a classic CoIP-MS assay to screen host proteins that interact with EspF in 293T cells. MS results identified 311 proteins. We then analyzed BPs, MFs, and CCs of the interacting proteins using GO enrichment, KEGG pathway, and STRING analyses.

Previous studies have identified that SNX9 can bind EPEC EspF via its amino-terminal SH3 region [31] and EspF can interact with host intermediate filament protein cytokeratin 18 (CK18) in a complex with 14-3-3 ε, which can alter the architecture of the intermediate filament network [32]. Moreover, EspF has the ability to interact with ribosomal proteins and modulate anti-phagocytosis by interacting with ANXA6 [33]. In our MS screening results, we also found that SNX9, SNX18, 14-3-3 ε, ANXA6, ANXA2 interacted with EspF, which confirmed our MS results.

MS results showed that host proteins interacting with EspF mostly involved in metabolic pathways and localized in the cytoplasm and nucleolus. We found that EspF interacted strongly with ribosomal RPL, RPS, and EIF family proteins, and can phosphorylate them. Previous proteomics studies have shown that many ribosomal protein levels in intestinal cells decrease after EPEC infection [34]. In cells expressing EspF, pre-rRNA synthesis is blocked. At the same time, EspF-dependent EPEC infection reduces the expression level of ribosomal protein RPL9, and changes the relocation of RPS5 and U8 small nucleolar RNA (snoRNA) [35]. Our results provided further support for the hypothesis that EspF may play its biological role by regulating ribosomal protein synthesis and relocation them. Future research will attempt to decipher the mechanism that EspF promotes inhibition of ribosome synthesis.

Meanwhile, we separately explored the interacting proteins and potential pathogenic mechanisms of the EspF-N/C terminus. We found that the proteins interacting with EspF-N terminus are mostly involved in metabolic pathways and protein processing in the endoplasmic reticulum. Our previous research proved that the N-terminus is essential for cell apoptosis, inflammatory response, and animal toxicity [36]. Thus, we speculate that EspF-N may also mediate apoptosis through endoplasmic reticulum stress. Previous studies have shown that the EspF-C terminus can bind to various host protein motifs. EPEC EspF interacts with SNX9 and N-WASP by the SH3 and CRIB motifs to induce actin polymerization and TJ disruption[17]. In our study, we found the host protein interacted with EspF-C-terminus mainly involved intermediate filaments

 doi:10.20944/preprints202012.0551.v1

and ribosomes. Among these proteins, TUBB showed strong interaction with EspF-C. We then hypothesed that EspF might cause cytoskeleton rearrangement by recruiting a multiple of host protein to the pedestal, including the interaction between its C-terminus and TUBB.

EspF targets mitochondria and regulates the expression of DNA mismatch repair proteins in host cells through post-transcriptional manipulation, leading to depletion of Apc and MMR proteins in host cells [37], and increases the instability of microsatellite DNA sites, which is a precursor to DNA damage and can even lead to colon cancer [38]. However, there is no research on whether EspF can directly cause DNA damage. Notably, we found that EspF can induce H2AX protein phosphorylation in Caco2 cells and induce cell multi-nuclearation and cell hypertrophy. We tried to find out which proteins EspF interacts with can mediate this pathogenic phenotype. We used CO-IP and IF to further validate the interaction between EspF and SMC1, and this interaction is mostly by EspF-C terminus. SMC1 is a chromosomal structural protein. When DNA is damaged, the upstream ATM will phosphorylate SMC1 to form the BRCA1-NBS1-MRE11 complex, and p-SMC1 activates the S-phase checkpoint in DNA damage repair [39]. Abnormal SMC1 expression or mutation can lead to a deficient DNA damage repair pathway, which is closely related to tumorigenesis [40]. Thus, we focused on the interaction between EspF and SMC1. We found that EspF can relocalize SMC1 to be more in the cytoplasm, thus we speculated EspF may also may reduce SMC1 protein in the nucleus, which decreases the cell's ability to repair DNA damage. EspF can also increase the expression of p-SMC1 both in $Caco_2$ and HT-29 cells. This series of evidence suggests that EspF may mediate DNA damage repair and cell proliferation by phosphorylating SMC1 (Figure 8). The precise mechanism is worthy of further study.

Furthermore, It is undeniable that CoIP-MS has certain limitations in the screening process. First, the bands for mass spectrometry should be distinct, and some weak interactions may not be detected. Second, different levels of protein expression between different cells may affect the binding efficiency of co-immunoprecipitation. We need further experiments to verify the real interactions between EspF and host cell proteins. In addition, EspF is not documented as a kinase, further study is needed on how it phosphorylates host proteins. But EspF can activate PKC α and phosphorylate it, which may mediate cell downstream protein phosphorylation.

In our research, we verified for the first time that EspF can cause DNA damage and EspF interacted with SMC1, most likely through its C-terminus. At the same time, EspF can increase p-SMC1 expression level and relocate SMC1 into the cytoplasm. This provides us with new insights into the role of EspF on mediating DNA damage repair. Further research is needed on what processes EspF is involved in and how these effects mediate bacterial–host pathogenesis. Our screening provides directions for future analysis of the potential biological role of EspF and its N and C terminus.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

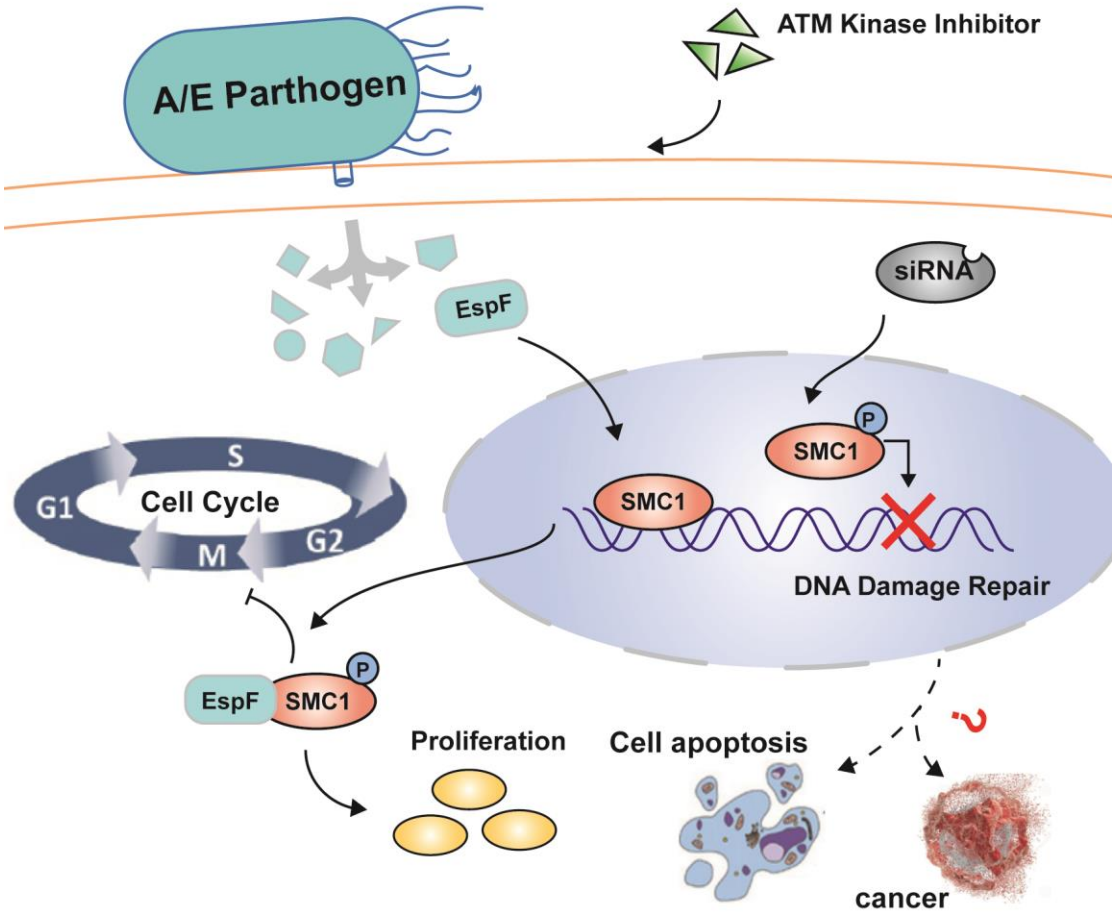



**Figure 8.** The biological effect of EspF by mediating on SMC1. EspF can increase the expression of p-SMC1 and relocalize it in the cytoplasm. Therefore, EspF may mediate DNA damage, cell proliferation, cell cycle arrest, cell apoptosis, and even cancer by interacting with SMC1.

## 4. Materials and Methods

### *4.1. Cell lines and strains*

293T and $Caco_2$ cells were cultured overnight in DMEM (Gibco) containing 10% fetal bovine serum (FBS) and 1% penicillin and streptomycin at 5% $CO_2$ and 37°C. HT-29 cells were cultured overnight in RPMI-1640 (Gibco) medium containing 10% FBS and 1% penicillin and streptomycin at 5% $CO_2$ and 37°C. The pEGFP-N1 plasmid, strain EHEC O157: H7 EDL 933, *ΔespF,* and *ΔespF/pespF* were stored in our laboratory [36]. DH5α competent cells and LA high-fidelity enzyme were purchased from TaKaRa. Restriction enzymes EcoR I and BamH I were purchased from Thermo Fisher Scientific. The primers used in this study were synthesized by Sangon Biotech (Shanghai, China). Gene sequencing was performed by IGE Biotechnology (Guangzhou, China).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

*4.2. Construction of pEGFP-EspF, pEGFP-EspF/N, and pEGFP-EspF/C-terminus plasmids*

We amplified the *espF* gene and its N-terminus (1–219 bp) and C-terminus (220–747 bp) using LA high-fidelity enzyme on the genome of EDL 933. Each constructed plasmid was labeled with Flag. Primers *espF-F*: CCGGAATTCGCCACCATGCTTAATGGAATTAGTAACGCTG

and *espF-R*: CGCGGATCCCCGCTACCGCCGCTTCCCTTGTCCTTATCGTCGTCATCCTTGTAATCCTTATCGTCGTCATCCTTGTAATCCTTATCGTCGTCATCCTTGTAATCCCCTTTCTTCGATTGCTCATA amplified the *espF* gene; primers *espF/N-F*: CCGGAATTCGCCACCATGCTTAATGGAATTAGACAC

and *espF / N-R*: CGCGGATCCCCACCAGAGCCACCCTTATCGTCGTCATCCTTGTAATCGGGAGTAAATGAAGTCACCTG amplified *espF/N*; and primers *espF/C-F*: CCGGAATTCGCCACCATGTCTCGTCCGGCACCGCGCGCCCCCACC and *espF/C-R*: CGCGGATCCCCACCTCCCC amplified the *espF/C*. The PCR product was digested with EcoRI/BamH I restriction enzyme and T-linked to the pEGFP-N1 plasmid to generate pEGFP-EspF, pEGFP-EspF/N, and pEGFP-EspF/C plasmids. Sequencing verified the integrity of the constructed plasmids.

293T cells were cultured according to standard methods, and the cells were plated on 10 $cm^2$ culture dishes (NEST, Hong Kong, China). In accordance with the manufacturer's instructions, the cells were transfected with pEGFP-EspF, pEGFP-EspF/N, and pEGFP-EspF/C plasmids using Lipofectamine 3000 (Thermo, USA).

*4.3. CoIP and western blot*

We followed the instructions of the co-immunoprecipitation Pierce CoIP Kit 26149 (Thermo, USA). 293T cells over-expressing pEGFP-EspF or pEGFP-EspF/N or pEGFP-EspF/C were prepared as mentioned above. 1 mg of the cell lysates were added 80 µl of the control agarose resin slurry to remove the non-specifically bound proteins. Next, we placed 20 µl of AminoLink resin slurry into the spin column, added 15 µg IgG antibody (Abclonal)/Flag antibody (CST), and co-incubated the mixture for 1 h. The protein was added to the spin column and incubated at 4°C overnight. The resin was washed three times with IP lysis, and then the Flag/IgG protein complex was eluted, Then underwent 10% SDS-PAGE electrophoresis for immunoblotting. The primary antibodies of rabbit anti-Flag (CST) and rabbit anti-SMC1 (CST) were added and incubated overnight at 4°C. After three washes with TBST, the corresponding secondary antibody was added, incubated at room temperature for 1 h, and the bands were visualized using an ECL chemical solution.

*4.4. Silver stain*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2020 doi:10.20944/preprints202012.0551.v1

After electrophoresis, the gels were incubated in a fixing solution for 15 min on a shaker at room temperature. The gel was transferred to sensitizer with gentle shaking. Then, we added the staining solution for 30 min on a shaker. After staining, the gel was washed with water three times (15 min each) and incubated in a developing solution until bands became visible. The gel was washed in water and then imaged. We then cut out the different bands and sent to Huijun Biological Company (Guangzhou, China) for mass spectrometry.

### *4.5. Bacterial infection*

EHEC O157: H7 EDL 933 was cultured in LB medium. *ΔespF* and *ΔespF/pespF* were cultured in LB broth medium containing kanamycin at 37°C for 12 hours, and HT-29/caco2 cells were seeded in 10 $cm^2$ culture dishes. When the cells reached 95% confluence, the cells were infected with bacteria at an MOI of 100: 1 and incubated in a incubator at 37°C and 5% $CO_2$ for 6 hours. Then, the medium was aspirated, the cells were gently washed with PBS, and the proteins were collected for immunoblotting.

### *4.6. Immunofluorescence assay*

Caco2 cells were plated on a confocal dish (35 mm, NEST, Hong) and transfected with pEGFP, pEGFP-EspF, pEGFP-EspF/N, and pEGFP-EspF/C vectors. After 48 hours, the cells were gently washed with PBS three times and fixed in 4% paraformaldehyde for 15 min at room temperature, then blocked with 0.1% Triton X-100 for 30 min. The fixed cells were stained with primary antibody overnight at 4°C(anti-p-H2AX/anti-CK18, CST), followed by incubation with the secondary antibody for 30 min, Then, the cells were stained with DAPI for 5 min. Cellular colocalization was observed under an FV1000 confocal microscope (Olympus, Tokyo, Japan).

## 5. Conclusions

In this study, we focused on 311 host proteins that interact with EHEC O157: H7 EspF and used bioinformatics enrichment to analyze their MFs, BPs, and cellular pathways. These findings provide new candidates for EspF interactions, suggesting that EspF can phosphorylate H2AX and may regulate DNA damage repair process by interacting with SMC1. These results are very encouraging. We also provided a PPI network of interactions between EspF/EspF-N/EspF-C terminus and the host, which can guide further research into the protein interactions that mediate the interplay between pathogen EHEC O157: H7 and host cells.

**Supplementary Materials:** Figure S1: GO enrichment, KEGG pathway, and STRING analyses of the 36 kDa proteins interacting with EspF. (A) GO annotation of identified EspF-interacting proteins in BPs, CCs, and MFs. (B) The KEGG analysis of

 doi:10.20944/preprints202012.0551.v1

the distribution of the top 20 protein interaction pathways. (C) STRING analysis of statistically significant proteins detected in proteins that interacted with EspF. Table S2: EspF-mediated phosphorylation, acetylation, and methylation of the host proteins.

**Author Contributions:** MQF designed the study; MQF, YH, and JLW conceived the methods; MQF, ZKZ, JLW, JYL, and XXL performed the experiments; and MQF, BZ, WZ and CSW discussed the data and wrote the paper.

**Funding:** This research was supported by the Natural Science Foundation of Guangdong Province (No.2018B030311063) and co-funded by No.2019A1515111004.

**Acknowledgments:** We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

**Conflicts of Interest:** The authors declare no conflflict of interest.

# References:

1. *DAVID Bioinformatics Resources 6.8.* In Vol. 295.
2. Al, J.M.R.E., Epidemiology of Escherichia coli O157:H7 Outbreaks, United States, 1982 – 2002 - Volume 11, Number 4—April 2005 - Emerging Infectious Disease journal - CDC. *EMERG INFECT DIS* **2005**, 11, (4), 603-609.
3. Eppinger, M.; Rasko, D.A.; Hazen, T.H.; Sadiq, S.M., EHEC Genomics: Past, Present, and Future. *Microbiology Spectrum* **2014**, 2, (4).
4. M, M.; P, P., Enterohemorrhagic Escherichia coli as the cause of diarrhea in the Czech Republic, 1965-2013. *Epidemiologie, mikrobiologie, imunologie : casopis Spolecnosti pro epidemiologii a mikrobiologii Ceske lekarske spolecnosti J.E. Purkyne* **2014**.
5. Weflen, A.W.; Alto, N.M.; Viswanathan, V.K.; Hecht, G., E.coli secreted protein F promotes EPEC invasion of intestinal epithelial cells via an SNX9-dependent mechanism. *CELL MICROBIOL* **2010**, 12, (7), 919-929.
6. Singh, A.P.; Sharma, S.; Pagarware, K.; Siraji, R.A.; Ansari, I.; Mandal, A.; Walling, P.; Aijaz, S., EnteropathogenicE. colieffectors EspF and Map independently disrupt tight junctions through distinct mechanisms involving transcriptional and post-transcriptional regulation. *SCI REP-UK* **2018**, 8, (1), 3719.
7. Zhao, S.Z.Y.W., The N-Terminal Domain of EspF Induces Host Cell Apoptosis after Infection with Enterohaemorrhagic Escherichia coliO157:H7. *PLOS ONE* **2010**, 8: e55164..
8. Charpentier, X.; Oswald, E., Identification of the secretion and translocation domain of the enteropathogenic and enterohemorrhagic Escherichia coli effector Cif, using TEM-1 beta-lactamase as a new fluorescence-based reporter. *J BACTERIOL* **2004**, 186, (16), 5486-5495.
9. Dean, P.; Scott, J.A.; Knox, A.A.; Quitard, S.; Watkins, N.J.; Kenny, B.; Van Nhieu, G.T., The Enteropathogenic E. coli Effector EspF Targets and Disrupts the Nucleolus by a Process Regulated by Mitochondrial Dysfunction. *PLOS PATHOG* **2010**, 6, (6), e1000961.
10. Roxas, J.L.; Koutsouris, A.; Viswanathan, V.K., Enteropathogenic Escherichia coli-Induced Epidermal Growth Factor Receptor Activation Contributes to Physiological Alterations in Intestinal Epithelial Cells. *INFECT IMMUN* **2007**, 75, (5), 2316-2324.

11. Holmes, A.; Muhlen, S.; Roe, A.J.; Dean, P., The EspF effector, a bacterial pathogen’ s swiss army knife. *Infection & Immunity* **2010**, 78, (11), 4445-4453.
12. Mcnamara, B.P.; Donnenberg, M.S., A novel proline-rich protein, EspF, is secreted from enteropathogenic *Escherichia coli* via the type III export pathway . *FEMS MICROBIOL LETT* **2006**, 166, (1), 71-78.
13. Meyuhas, O., Ribosomal Protein S6 Phosphorylation: Four Decades of Research. *International Review of Cell & Molecular Biology* **2015**, 320, 41.
14. Enerly, E.; AHMADI, H.; SHALCHIAN-TABRIZI, K.; LAMBERTSSON, A., Identification and comparative analysis of the RpL14 gene from Takifugu rubripes. *HEREDITAS* **2003**, 139, (2), 143-150.
15. Boye, E.; Grallert, B., eIF2α phosphorylation and the regulation of translation. *CURR GENET* **2019**, (84–98).
16. Ti, S.; Alushin, G.M.; Kapoor, T.M., Human β-Tubulin Isotypes Can Regulate Microtubule Protofilament Number and Stability. *DEV CELL* **2018**.
17. Marches, O.; Batchelor, M.; Shaw, R.K.; Patel, A.; Cummings, N.; Nagai, T.; Sasakawa, C.; Carlsson, S.R.; Lundmark, R.; Cougoule, C., EspF of Enteropathogenic Escherichia coli Binds Sorting Nexin 9. *J BACTERIOL* **2006**, 188, (8), 3110-3115.
18. Garber, J.J.; Mallick, E.M.; Scanlon, K.; Turner, J.R.; Donnenberg, M.S.; Leong, J.M.; Snapper, S.B., Attaching-and-Effacing Pathogens Exploit Junction Regulatory Activities of N-WASP and SNX9 to Disrupt the Intestinal Barrier. **2018**.
19. Hua, Y.; Ju, J.; Wang, X.; Zhang, B.; Zhao, W.; Zhang, Q.; Feng, Y.; Ma, W.; Wan, C., Screening for host proteins interacting withEscherichia coli O157:H7 EspF using bimolecular fluorescence complementation. *FUTURE MICROBIOL* **2018**, 13, (1), 37-58.
20. Díaz-Díaz, A.; Roca-Lema, D.; Casas-Pais, A.; Romay, G.; Colombo, G.; Concha, Á.; Graña, B.; Figueroa, A., Heat Shock Protein 90 Chaperone Regulates the E3 Ubiquitin-Ligase Hakai Protein Stability. *CANCERS* **2020**, 12, (1), 215.
21. Li, J.; He, J.; Wang, Y.; Shu, Y.; Zhou, J., SMC1 promotes proliferation and inhibits apoptosis through the NF-κB signaling pathway in colorectal cancer. *ONCOL REP* **2019**, 42, (4), 1329-1342.
22. Lidza; Kalifa; Jennifer; S.; Gewandter; Rhonda; J.; Staversky; Elaine; A., Kalifa L , Gewandter J S , Staversky R J , et al. DNA double-strand breaks activate ATM independent of mitochondrial dysfunction in A549 cells. *FREE RADICAL BIO MED* **2014**, (75:30-39).
23. Zhang, Y.; Zhang, X.; Fleming, M.R.; Amiri, A.; El-Hassar, L.; Surguchev, A.A.; Hyland, C.; Jenkins, D.P.; Desai, R.; Brown, M.R., Kv3.3 Channels Bind Hax-1 and Arp2/3 to Assemble a Stable Local Actin Network that Regulates Channel Gating. *CELL* **2016**, S0092867416301118.
24. Jefferson, M.; Bone, B.; Buck, J.L.; Powell, P.P., The Autophagy Protein ATG16L1 Is Required for Sindbis Virus-Induced eIF2α Phosphorylation and Stress Granule Formation. *Viruses* **2020**, 12, (1), 39.
25. Nougayrède, J.P.; Donnenberg, M.S., Enteropathogenic Escherichia coli EspF is targeted to mitochondria and is required to initiate the mitochondrial death pathway. *CELL MICROBIOL* **2004**, 6, (11), 1097-1111.
26. Pardal, A.J.; Fernandes-Duarte, F.; Bowman, A.J., The histone chaperoning pathway: from ribosome to nucleosome. *ESSAYS BIOCHEM* **2019**, 63, (1), 29-43.
27. Nagelkerke, A.; Span, P.N., *Staining Against Phospho-H2AX ( γ-H2AX) as a Marker for DNA Damage and Genomic Instability in Cancer Tissues and Cells*. Springer International Publishing,

2016.

28. Dean, P.; Kenny, B., A bacterial encoded protein induces extreme multinucleation and cell-cell internalization in intestinal cells. *Tissue Barriers* **2013**, 1, (1), e22639.
29. Viswanathan, V.K.; Lukic, S.; Koutsouris, A.; Miao, R.; Muza, M.M.; Hecht, G., Cytokeratin 18 interacts with the enteropathogenic Escherichia coli secreted protein F (EspF) and is redistributed after infection. *CELL MICROBIOL* **2004**, 6, (10), 987-997.
30. Nougayrède, J.P.; Foster, G.H.; Donnenberg, M.S., Enteropathogenic Escherichia coli effector EspF interacts with host protein Abcf2. *CELL MICROBIOL* **2007**, 9, (3), 680-693.
31. Oliver, M.; Miranda, B.; Robert K, S.; Amit, P.; Nicola, C.; Takeshi, N.; Chihiro, S.; Sven R, C.; Richard, L.; Celine, C.; Emmanuelle, C.; Stuart, K.; Ian, C.; Gad, F., EspF of enteropathogenic Escherichia coli binds sorting nexin 9. *J BACTERIOL* **2006**, 188, (8).
32. Viswanathan, V.K.; Lukic, S.; Koutsouris, A.; Miao, R.; Hecht, G., Cytokeratin 18 interacts with the enteropathogenic Escherichia coli secreted protein F (EspF) and is redistributed after infection. *CELL MICROBIOL* **2010**, 6, (10), 987-997.
33. Hua, Y.; Ju, J.; Wang, X.; Zhang, B.; Zhao, W.; Zhang, Q.; Feng, Y.; Ma, W.; Wan, C., Screening for host proteins interacting withEscherichia coli O157:H7 EspF using bimolecular fluorescence complementation. *FUTURE MICROBIOL* **2018**, 13, (1), 37-58.
34. Hardwidge, P.R.; Rodriguez-Escudero, I.; Goode, D.; Donohoe, S.; Eng, J.; Goodlett, D.R.; Aebersold, R.; Finlay, B.B., Proteomic Analysis of the Intestinal Epithelial Cell Response to Enteropathogenic Escherichia coli. *J BIOL CHEM* **2004**, 279, (19), 20127-20136.
35. Dean, P.; Scott, J.A.; Knox, A.A.; Quitard, S.; Watkins, N.J.; Kenny, B., The enteropathogenic E. coli effector EspF targets and disrupts the nucleolus by a process regulated by mitochondrial dysfunction. *PLOS PATHOG* **2010**, 6, (6), e1000961.
36. Wang, X.; Du, Y.; Hua, Y.; Fu, M.; Niu, C.; Zhang, B.; Zhao, W.; Zhang, Q.; Wan, C., The EspF N-Terminal of Enterohemorrhagic Escherichia coli O157:H7 EDL933w Imparts Stronger Toxicity Effects on HT-29 Cells than the C-Terminal. *FRONT CELL INFECT MI* **2017**, 7.
37. Maddocks, O.D.K.; Short, A.J.; Donnenberg, M.S.; Bader, S.; Harrison, D.J., Attaching and Effacing Escherichia coli Downregulate DNA Mismatch Repair Protein In Vitro and Are Associated with Colorectal Adenocarcinomas in Humans. *PLOS ONE* **2009**, 4.
38. Maddocks, O.D.K.; Scanlon, K.M.; Donnenberg, M.S., An Escherichia coli Effector Protein Promotes Host Mutation via Depletion of DNA Mismatch Repair Proteins. *MBIO* **2013**, 4, (3), e00152-13.
39. Luo, Y.; Deng, X.; Cheng, F.; Li, Y.; Qiu, J., SMC1-Mediated Intra-S-Phase Arrest Facilitates Bocavirus DNA Replication. *J VIROL* **2013**, 87, (7), 4017-4032.
40. Ying, Z.; Fei, Y.; Liang, W.; Zhuo, W.; Naijin, Z.; Zhijun, W.; Ziwei, L.; Xiaoyu, S.; Shi, W.; Liu, C., Phosphorylation of SMC1A promotes hepatocellular carcinoma cell proliferation and migration. *INT J BIOL SCI* **2018**, 14, (9), 1081-1089.