Iranian Journal of Blood & Cancer

Iran J Blood Cancer, 2024, 16(3), 63-69
https://ijbc.ir

Review

The Positive Impact of Virulence Genes in Bacterial Pathogenicity and Colonization on the Large Bowel in Colorectal Cancer Patients

Mohammad Khalilollahi [1], Fatemeh Ashrafi [1], Arash Moradi [2], Jean-Philippe Nougayrède [3], Shahla Mohammad Ganji [2,*]

[1] *Department of Microbiology, Islamic Azad University, North Tehran Branch, Tehran, Iran.*
[2] *Departments of Molecular Medicine, National Institute of Genetic Engineering and Biotechnology (NIGEB), Tehran, Iran.*
[3] *INRA, UMR1225, F-31076 Toulouse, France, and Universite´ de Toulouse, ENVT, UMR1225, F-31076 Toulouse, France.*

Scan and read the article online

Citation Mohammad Khalilollahi, Fatemeh Ashrafi, Arash Moradi, Jean-Philippe Nougayrède, Shahla Mohammad Ganji. The Positive impact of virulence genes in bacterial pathogenicity and colonization on the large bowel in colorectal cancer patients. Iran J Blood Cancer. 2024 Sep 30;16(3): 63-69.

doi 0

Article info:
Received: 19 Jun 2024
Accepted: 03 Sep 2024
Published: 30 Sep 2024

Keywords:
Escherichia coli (*E. coli*)
Intestine's inflammatory disease (IBD)
Virulence genes
Colorectal cancer (CRC)

Abstract

Colorectal Cancer (CRC) is the third most common cancer around the world. Among the various factors associated with the development of CRC, bacterial infection and related toxins are considered the most critical risk factors. Several virulence genes, including fimC, cnf1, vat1, hlyD, clbB, clbN, feoA, feoB, fyuA, iroN, ireA, iutA, KpsMT (k1), KpsMTII, and KpsMTIII, have been found to have a more significant influence on the pathogenicity of the bacteria. This research aimed to assess the possible position performed through a few virulence genes in E. coli isolated from the intestinal tissues' biopsies of patients with colorectal cancer. Using microbial and biochemical methods, this study isolated 82 samples of E. coli from all of the 170 biopsies obtained from patients suffering from CRC, inflammatory bowel disease, and normal individuals. Then, the frequency of 15 virulence genes was assessed by applying PCR. The obtained results indicated that two types of bacterial genes as following are more likely to be involved in CRC development: clbB and clbN genes, which are associated with the colibactin polyketide synthesis system, as well as KpsMTIII gene, which is involved in polysaccharide capsule synthesis. In precis, these consequences suggest that the superiority of E. coli containing clbB, clbN, and KpsMTIII plays an extensive role in the inflammation and, therefore, the occurrence of CRC.

1. INTRODUCTION

The global burden of colorectal cancer (CRC) has become a significant public health concern, including in Iran. Various risk factors, such as red or processed meat consumption, alcohol, and chronic digestive tract inflammation, have been linked to CRC incidence (1).

Recent studies established an association between inflammatory bowel disease (IBD) and CRC (2). The microbiome plays a critical role in developing CRC, with alterations in its composition and function potentially contributing to conditions like IBD and CRC (3).

* **Corresponding Author:**
Shahla Mohammad Ganji
Affiliation: *Departments of Molecular Medicine, National Institute of Genetic Engineering and Biotechnology (NIGEB), Tehran, Iran.*
E-mail: **shahla@nigeb.ac.ir**

[DOI: 10.61186/ijbc.16.3.63]

The *E. coli* species of the phylogenetic group B2 have a genomic island called *PKS*, also called *E. coli* "*PKS+*" species. This genomic island, "*PKS+*," codes the production of a polyketide-peptide genotoxin, colibactin, and induces an increased mutation rate (4). A high percentage of mucosa-associated *PKS*-positive *E. coli* could be found in patients with CRC and IBD. The experiments have shown that *E. coli* specimens with *PKS+* can lead to DNA damage in both human and animal epithelial cells. The *E. coli* strains that promote the development of CRC need more virulence genes (5). The characteristic virulence elements of *E. coli* are genetically coded by using chromosomal, plasmid, and bacteriophage DNAs, respectively.

Several bacterial properties, which include P fimbriae, fimbriae kind 1, hemolysin, aerobactin, serum resistance, and the Kl capsule, are nicely established as virulence factors in the intense symptomatic *E. coli* pressure. The virulence factors related genes include the following set of genes: *fim*C, *cnf1*, *vat1*, *hlyD*, *clbB* and *clbN*, *feo*A, *feoB*, *fyu*A, *iroN*, *ire*A, iutA, *KpsMT (k1)*, *KpsMTII*, and *KpsMTIII* which have an essential role in the pathogenesis of the bacteria (6). As far as we know, no former reports indicate an investigation into the role that *E. coli* virulence factors might play in the induction and progression of the CRC in Iran. This report evaluates and identifies the virulence genes of the *E. coli* isolated from the intestinal tissue biopsy based on their position as the inner membrane, such *as holy*, *fim*C, *cnf1*, *vat1* (7), and outer membrane. The pathogenesis factors associated with the outer membrane iron receptor genes are *fyuA*, *feoA*, *feoB*, *ireA*, *iroN*, *iutA* [aerobatic system], and polysaccharide capsule marker genes are *kpsMT (K1)*, *kpsMTII* and *kpsMT III*.

This study investigates the impact of the virulence genes in the *E. coli* pathogenicity and colonization in the large bowel of patients with colorectal cancers. To achieve this goal, we have studied the prevalence of the 15 virulence genes in E. coli isolated from tissue specimens of individuals with CRC, IBD, and normal individuals.

2. MATERIALS AND METHODS

2.1. Sample collection

One hundred and seventy tissue samples were collected from Shahid Beheshti Hospital in Qom and the Tumor Bank Center at Imam Khomeini Hospital in Tehran from June 2014 to February 2015. Following surgical removal of the specimens, tissues were immediately transferred to the laboratory, and pathological information was collected following pathological examination by the pathology department of either of the hospitals from which specimens were obtained. The collected samples were taken from subjects suffering from colorectal cancer (60 cases), the inflammatory disease of the colon (*i.e.*, large intestine inflammatory disorders, polyps) (60 cases), and also from 50 normal subjects participating in this study. All participants filled out the questionnaires and consent forms.

2.2. Bacterial isolation and assessment

The fresh biopsies obtained from hospitals were incubated at 37 °C in an Erlenmeyer flask containing LB [Luria Bertoni broth] for 24 hours. Then, serial dilutions were prepared in tubes containing the physiological saline solution. Next, each tube containing serially diluted bacterial solution was cultured in the LB agar plates, and the grown bacterial colonies were transferred onto EMB [Eosin-methylene blue] agar medium. The green sheen colonies obtained in this environment underwent the biochemical IMVIC analysis of indole, methyl red (MR), citrate, Voges-Proskauer (VP), and citrate utilization tests, respectively. The typical specimens that were confirmed to be *E. coli* were used for further study or stored at -20 °C. Following the microbial and biochemical analysis, 82 specimens of *E. coli* have been removed from the 170 intestinal biopsies. As detailed, 7 *E. coli* (14%) were remoted from ordinary sufferers, 46 (76.6%) *E. coli* were from CRC patients, and 29 *E. coli* (48.3%) were isolated from patients with IBD.

2.3. Molecular analysis

DNA Extraction and Purification: The DNA was extracted from all isolated bacteria and purified using the genome DNA purification kit. After DNA extraction, the specimens were evaluated for the presence of the virulence genes, including *fim*C, *cnf1*, *vat1*, *hlyD*, *clbB* and *clbN*, *feo*A, *feoB*, *fyu*A, *iroN*, *ire*A, *iut*A, *KpsMT (k1)*, *KpsMT II*, and *KpsMT III*, using polymerase chain reaction [PCR]. The specific primers designed by present scientific group and used for the detection of the specific sequence of the genes [i.e., *fim*C, *cnf1*, *vat1*, *hlyD*, *clbB and clbN*, *feo*A, *feoB*, *fyu*A, *iroN*, *ire*A, *iut*A, *KpsMT (k1)*, *KpsMT II*, and *KpsMT III*] are genes that are involved in the biosynthesis of the *E. coli* virulence factors. We used online and offline databases such as NCBI, BLAST, and OLIGO7 software for primer designing. The positive control of *E. coli* with *PKS* island genes was gifted by Prof. Jean-Philippe Nougayrede from the University of Toulouse, France. An amount of 10µl master mix (Ampliqon, Dk.), 0.5µl of each F and R primers with a concentration of 10 pmol, 1µl of extracted DNA template with a concentration of 100 ng, and 8µl of distilled water

[DOI: 10.61186/ijbc.16.3.63]

were mixed in 0.5 ml PCR microtubes and were mixed well. Finally, the microtubes were placed in an Eppendorf thermocycler. The reaction was started with an initial denaturation of 2 minutes at a temperature of 93 °C, followed by 35 cycles of denaturation at 93 °C for 35 seconds, an annealing time of 30 seconds, and an extension at 72 °C for 35 seconds. Ultimately, the temperature was set to 72 °C for 10 minutes. The sequences of primers and annealing temperature for selected genes are reported in **Table 1**. The PCR products were electrophoresed on a 1% agarose gel. The amplified DNA was stored simultaneously using ethidium bromide-containing loading dye, and gels were analyzed using gel documentation.

2.4. Molecular analysis

The phylogenetic tree was constructed for the *clbB* and *clbN*-positive isolates of the *E. coli* strains based on 16srRNA sequencing. The sequencing of the 16S rRNA gene was conducted using a genetic analyzer (ABI Prism 3130 and 3130 xl Genetic Analyzer) at Pishgam Co., Korea, for the *clbB* and *clbN* positive isolates of the *E. coli* strains. The sequencing readings were done as paired-end. The sequences were edited to exclude the PCR primer binding sites and manually corrected using DNA Star, BioEdit software, and MEGA6 version software. The full gene sequences of the strains were compared automatically using BLAST against the sequences of bacteria available in databanks (http://www.ncbi.nlm.nih.gov/). The phylogenetic analysis was constructed using a neighbor-joining algorithm. The complete sequences (1380 bp) of the 16S rRNA gene of *E. coli* strains NGR01, NGR04, NGR05, NGR06, and Control Positive, as gifted, have been deposited in the International Nucleotide Sequence Database (INSD), that is, in the National Center for Biotechnology Information (NCBI). Their accession numbers are KM016094.1, KM016093.1, KM016092.1, KM016091.1, KM016090.1, KM016089.1, KM016088.1, KJ801975.1, and KM016095.1.

2.5. Statistical Analysis

Kruskal-Wallis nonparametric analysis of variance (ANOVA) was used to compare the correlation of virulence genes in isolated *E. coli* bacteria from the tissues of individuals with CRC, IBD, and normal individuals. The Fisher's exact test (two-tailed) was used to compare the proportions of patient samples positive for *E. coli* across the above populations. Unpaired Student t-tests and Mann-Whitney U tests were used to compare two groups where appropriate. The P value of <0.05 was considered statistically significant. Statistical calculations were performed using the Statistical Package for the Social Sciences software (SPSS, Chicago, IL) 16.0.

3. RESULT

3.1. Microbial confirmation finding

The results of the IMVIC test for *E. coli* bacteria for the cases of indole (I) and methyl red (MR) were positive, while those for VP and citrate were negative. Therefore, isolating and carrying out the typing and subsequent experiments on the isolated bacteria was possible. In this study, 82 biopsies were analyzed. Among this total number of analyzed biopsies, 7 out of 50 (14%) were specimens obtained from healthy participating individuals, 46 out of 60 (76.6%) were patients with colorectal cancer, and 29 out of 60 (48.3%) were patients with other colorectal inflammatory diseases of the colon polyps, as summarized in **Table 2**. Due to several limitations for sampling, some patients did not complete the questionnaire fully; hence, their relevant samples were excluded for further analysis. The obtained results showed a significant ($p<0.05$) correlation between inflammatory disease of the colon, colorectal cancer, and bacterial infection. The significant correlation parameters were weight loss in CRC versus standard cases (P value =0.004) and IBD versus normal (P value =0.04). The results indicate no significant difference between the studied groups regarding gender, height, smoking, consumption of meat and vegetables per week, taking multivitamins or aspirin, and family history.

3.2. Molecular finding

The molecular analysis of the selected virulence genes on the isolated *E. coli* from 82 samples from IBD, CRC, and normal individuals showed a significant difference between the studied participants in KpsMT III, clbB, and clbN. The results showed no significant difference between other virulence genes studies and the three types of studied samples (**Table 2**).

As detailed, the fragment length for clbB and *clbN* genes on the electrophoresis gel was 500 and 700 bp, respectively. The frequency of presence of the *clbB* and clbN genes isolated from tissue samples of patients with CRC, IBD, and the normal group were 71.7%, 51.7%, and 28.6%, respectively, with significant differences ($p = 0.04$). The frequency of the *KpsMT III* between these three groups was 100%, 75.8%, and 14.28% respectively (**Table 2**).

A study on the association between *clbB* or *clbN* with the other bacterial traits regarding CRC cancer shows that among the 82 isolated samples of *E. coli*, there could be a

[DOI: 10.61186/ijbc.16.3.63]

Table 1. The list of the specific primers used to detect the specific sequence of the virulence genes.

Gene name	Nucleotide Sequence (5'- 3')	Size of PCR product (bp)	Annealing Temp (^{O}C)
fimC	F: GTTTTATCGTGACGCCACCT R: TTCCTGCATCAGAAGGCAAT	408	48 ^{O}C
hlyD	F: CTCCGGTACGTGAAAAGGAC R: GCCCTGATTACTGAAGCCTG	320	48.5 ^{O}C
cnf1	F: CTTTACAATATTGACATGCTG R: TCGTTATAAAATCAAACAGTG	1105	40.5 ^{O}C
Vat1	F: GTGTCAGAACGGAATTGTC R: GGGTATCTGTATCATGGCAAG	230	47.7 ^{O}C
feoB	F: AAGTCAAAGCAGGGGTTGCGGG R: GACGCCGACATTAAGACGCGC	390	62 ^{O}C
FyuA	F: AGGGGGCACAACTGATTCCGC R: TACCGGGCCGTTTTCTGCCG	190	54 ^{O}C
ClbB	F: GATTTGGATACTGGCGATAACCG R: CCATTTCCCGTTTGAGCACAC	500	56 ^{O}C
ClbN	F: GTTTTGCTCGCCAGATAGTCATTC R: CAGTTCGGGTATGTGTGGAAGG	700	54^{O}C
FeoA	F: GTCAAAGGGTTAAGCAGGCGGG R: CATTAAGACGCGCGACGCCGA	850	58 ^{O}C
iroN	F: AACAGGGGGGATCACTTCCGC R: TACCGTTCCGGGTTCTGGCC	230	56 ^{O}C
ireA	F: CGCGCGGGATCCTCTGATAAAAAAGAAGAT R: ATATATAAGCTTGAAGGATACTCTTACATT	460	55^{O}C
iutA	F: CCCGACGGTAAATGCGAATAAACAGG R: CAGCATCCAGCCCCCAGGTGA	210	62 ^{O}C
KpsMT (k1)	F: TAGCAAACGTTCTATTGGTGC R: CATCCAGACGATAAGCATGAGCA	395	60^{O}C
KpsMTII	F: GCGCATTTGCTGATACTGTTG R: CATCCAGACGATAAGCATGAGCA	380	56 ^{O}C
KpsMTIII	F: TCCTCTTGCTACTATTCCCCCT R: AGGCGTATCCATCCCTCCTAAC	240	49 ^{O}C

strong association between these two genes and several other virulence genes studied herein, including *fimC*, *vat1*, *cnf1*, *hlyD*, *iroN*, *fyuA*, *feoA*, *feoB*, *ireA*, *iutA*, *KpsMT (K1)*, *KpsMTII*, and *KpsMTIII* in Iranian patients (**Table 3**). These findings suggest that colibactin possibly contributes to the enhanced virulence and susceptibility of CRC cancer. The results from the constructed phylogenetic tree confirm that the isolated bacteria are *E. coli* strain *PKS* positive. The phylogenetic tree also shows that the *E. coli* specimens isolated from tissues were positive for the two *PKS* island genes, *clbB* and *clbN*. The results showed that *E. coli* bacteria isolated in this study are highly similar to the positive control explained above. The observed genetic distance from the other bacteria in this tree indicates that they belong to different groups of *E. coli* (**Figure 1, supplementary section**).

4. DISCUSSION

Recent investigations have identified a significant association between the composition of the gut microbiome and the pathogenesis of specific cancers (8). Bonnet et al. (7), have elucidated a potential mechanism, positing that pathogenic strains of *Escherichia coli* can serve as cofactors in colorectal cancer development. These *E. coli* strains may contribute to an elevated mutation rate (9). The virulence of *E. coli* is likely influenced by a complex interplay of factors, including invasion capabilities, toxin production, acquisition systems, adhesion properties, and colonization potential.

This study investigated the prevalence of 15 virulence genes in 82 *E. coli* bacterial isolates obtained from 170 biopsy specimens of CRC patients, IBD patients, and a normal control group. The findings of this research, when compared with previous studies, suggest a potential association between these virulence genes and CRC. This association may be mediated through the induction of inflammation and increased mutability within *E. coli*. A former study showed a statistical association between the *PKS* Island and the sequences involved in adherence, iron acquisition, and lipopolysaccharide synthesis (10). Ultimately, this molecular study suggests that *clbB*, *clbN* ($p=0.04$), and *KpsMT III* ($p< 0.001$) genes play an essential role in IBD disease and colorectal cancer. Furthermore, Johnson et al. 2008 studied the phylogenetic and molecular epidemiology distribution of the *E. coli PKS* genomic island (11). Moreover, they conveyed similarities in *clbB* and *clbN*

[DOI: 10.61186/ijbc.16.3.63]

Table 2. The comparison of correlation of virulence genes in isolated E. coli bacteria from tissues of individuals with CRC, IBD, and Normal individuals.

Category and trait	Total (n=170) %	Normal (n=50) %	CRC (n=60) %	IBD (n=60) %	*P*value	Odds ratio	95% confidence interval	R^{2d}
Isolated *E. coli*	82 (48.23%)	7 (14%)	46 (76.6%)	29 (48.3%)	-	-	-	-
Adhesin								
fimC	72 (87.8%)	5 (71.4%)	43 (93.5%)	24 (82.8%)	0.201	0.411	0.107-1.582	0.027
Toxins								
hlyD	36 (43.9%)	3 (42.9%)	22 (48.9%)	14 (48.3%)	0.76	0.982	0.415-2.323	< 0.001
cnf1	36 (43.9%)	3 (42.9%)	22 (47.9%)	14 (48.3%)	0.76	0.947	0.400-2.245	0.005
vat	66 (80.5%)	7 (100%)	39 (84.8%)	25 (86.2%)	0.47	1.23	0.332-4.577	0.002
clbB and *clbN*	50 (60.9%)	2 (28.6%)	33 (71.73%)	15 (51.72%)	0.04*	0.629	0.263-1.505	0.017
Siderophores								
iroN	32 (39.0%)	0	21 (45.6%)	11 (37.9%)	0.15	0.995	0.409-2.424	< 0.001
fyuA	13 (15.8%)	1 (14.28%)	8 (17.4%)	4 (13.8%)	0.9	0.602	0.170-2.135	0.010
feoA	43 (52.4%)	2 (28.6%)	25 (54.3%)	16 (55.2%)	0.47	1.025	0.432-2.432	< 0.001
feoB	45 (54.9%)	3 (42.6%)	27 (58.7%)	15 (51.72%)	0.71	0.776	0.328-1.838	0.005
ireA	21 (25.6%)	1 (14.28%)	12 (26.1%)	8 (27.6%)	0.86	0.783	0.288-2.129	0.004
iutA	23 (28.0%)	2 (28.6%)	14 (30.4%)	7 (24.1%)	0.74	0.547	0.198-1.513	0.022
Capsule								
KpsMT(K1)	77 (93.9%)	2 (28.6%)	46 (100%)	29 (100%)	0.76	~~	~~~	< 0.001
KpsMTII	73 (89.0%)	1 (14.28%)	45 (97.8%)	27 (93.1%)	0.90	2.286	0.572-9.133	0.023
KpsMTIII	70 (85.4%)	1 (14.28%)	46 (100%)	22 (75.8%)	< 0.001*	5.645	1.186-26.863	0.096

(p= 0.05) genes in *E. coli* isolated from hospitalized veterans (62 blood and fecal isolates and 69 fecal isolates). However, in this study, few specimens contained the two genes simultaneously. In addition, their result indicated that *clbB* and *clbN* are significantly associated with many other virulence genes. Our study is consistent with the results of Johnson et al. (11).

Our study found that among the 15 newly studied virulence genes, only *KpsMTIII*, *clbB*, and *clbN* were significantly associated with colorectal cancer in statistical analyses. Previous research has highlighted the importance of bacterial adherence to the host as a critical step in colonization and the success of microbial pathogens (12). The ability to adhere allows bacteria to start an infection successfully. Colonization of *E. coli* is accomplished by fimbriae or pili. It was proposed that establishing *E. coli* in the host intestinal tissue may help the successful development of colon cancer (13). The high-frequency *fimC* virulence gene can be considered evidence for this proposal (14). The presence of virulent *E. coli* specimens in the mucosa of CRC patients indicates their role in carcinogenesis (15).

According to Bonnet et al. findings in 2014, not only the mucosa-adherent *E. coli* not limited to the tumor site but also the total number of *E. coli* in the intestinal environment has a role in developing CRC (7). It can be argued that environmental changes such as loss of mucosal barrier and polarity of the epithelial cells' surface antigens could also augment the binding affinity of the *E. coli* to the epithelial cells and integration of the *E. coli* in the intestinal tissue (16). Raisch et al. (17) showed that tumor growth could be observed in mice injected with positive bacteria for *PKS* and a primary stimulus (e.g., a mutated APC). The results of this study confirm the presence of virulence genes in the three groups, including the isolated biopsy from normal subjects, patients with IBD, and patients with CRC. These results align with other research by other researchers in this field. Several virulence factors are attributed to the bacterium *E. coli* pathogenicity. A better understanding of the pathogenic organisms' virulence properties allows physicians to follow, treat, and prevent infection (18).

In conclusion, this study shows that *E. coli* colibactin synthesis genes, including *clbB*, *clbN*, and *KpsMTIII*, are significantly associated with inflammation and an increased mutation rate of the other virulence genes. These findings confirm and extend the previous works regarding *clbB* and *clbN* and support investigating the colibactin system and polysaccharide capsule marker *KpsMTIII* genes as a potential target for preventive or therapeutic measures. The presence

[DOI: 10.61186/ijbc.16.3.63]

Table 3. The Association of the clbB or clbN with the other bacterial traits among the 82 isolates of Escherichia coli isolates from Iranian patients with colorectal cancer.

Category and trait	Prevalence of trait according to *clbB* or *clbN* status (no.%)		*P*-value
	clbB and *clbN* negative (*n*=32)	*clbB* and *clbN* positive (*n=50*)	
Adhesins			
fimC	10 (12.19%)	72 (87.80%)	< 0.001*
Toxins			
hlyD	46 (56.9%)	36 (43.90%)	0.04*
cnf1	46 (56.9%)	36 (43.90%)	0.04*
vat	16 (19.51%)	66 (80.49%)	0.01*
Siderophores			
iroN	50 (60.97%)	32 (30.02%)	< 0.001*
fyuA	69 (84.14%)	13 (15.85%)	< 0.001*
feoA	39 (47.56%)	43 (54.43%)	0.37*
feoB	61 (74.39%)	21 (25.60%)	< 0.001*
ireA	59 (71.95%)	23 (28.05%)	< 0.001*
iutA			
Capsule			
KpsMT(K1)	5 (6.09%)	77 (93.90%)	< 0.001*
KpsMTII	9 (10.97%)	73 (89.02%)	< 0.001*
KpsMT III	12 (14.63)	70 (85.37)	< 0.001*

of virulence genes in the intestinal flora could predispose to developing several types of cancer, including colorectal cancer. In addition, the virulence genes in these bacteria were confirmed with a significant frequency. Thus, by understanding the critical role of the genes in isolated specimens in this study that cause intestinal infections, inflammation, and colorectal cancer, we could prevent these diseases or provide treatment strategies for healthcare practitioners.

Acknowledgment

The authors would like to thank the Iran Tumor Bank, the Cancer Institute of Iran in the Imam Khomeini Hospital Complex, and Shahid Beheshti Hospital in Qom for their kind collaboration and help in providing the project with tissue samples.

Conflict of interest

The authors declare no conflict of interest.

Funding

The present study was partly supported by postgraduate students and partly supported by the National Institute of Genetic Engineering and Biotechnology (NIGEB) using advanced equipment. The funding bodies had no role in the study design, data collection, analysis, interpretation of data, or manuscript preparation.

Ethical statement

The Ethics Committee approved the experimental protocols of the National Institute of Genetic Engineering and Biotechnology. For the part of the research where patient samples were obtained from the hospital, the code of ethics was taken, which is IR.NIGEB.EC.1394.12.01.G. The rest of the laboratory work was also done on bacteria. Written informed consent was obtained from individual or guardian participants. All methods were carried out following relevant guidelines and regulations.

Availability of data and materials:

Raw sequencing data have been deposited in the National Center for Biotechnology Information (NCBI), GenBank website (www.ncbi.nlm.nih.gov/genbank), and have the accession numbers KM016094.1, KM016093.1, KM016092.1, KM016091.1, KM016090.1, KM016089.1, KM016088.1, KJ801975.1, KM016095.1.

References

1. Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. 2023. Colorectal cancer statistics, 2023. *CA Cancer J Clin*.73(3):233-54.

2. Sato Y, Tsujinaka S, Miura T, Kitamura Y, Suzuki H, Shibata C. 2023. Inflammatory Bowel Disease and Colorectal Cancer: Epidemiology, Etiology, Surveillance, and Management. *Cancers (Basel)*.15(16).

3. Grellier N, Severino A, Archilei S, Kim J, Gasbarrini A, Cammarota G, et al. 2024. Gut microbiota in inflammation and

[DOI: 10.61186/ijbc.16.3.63]

colorectal cancer: A potential Toolbox for Clinicians. *Best Practice & Research Clinical Gastroenterology*.101942.

4. Iyadorai T, Mariappan V, Vellasamy KM, Wanyiri JW, Roslani AC, Lee GK, et al. 2020. Prevalence and association of pks+ Escherichia coli with colorectal cancer in patients at the University Malaya Medical Centre, Malaysia. *PLoS One*.15(1):e0228217.

5. Arthur JC, Perez-Chanona E, Mühlbauer M, Tomkovich S, Uronis JM, Fan TJ, et al. 2012. Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science*.338(6103):120-3.

6. Sarowska J, Futoma-Koloch B, Jama-Kmiecik A, Frej-Madrzak M, Ksiazczyk M, Bugla-Ploskonska G, Choroszy-Krol I. 2019. Virulence factors, prevalence and potential transmission of extraintestinal pathogenic Escherichia coli isolated from different sources: recent reports. *Gut Pathog*.11:10.

7. Bonnet M, Buc E, Sauvanet P, Darcha C, Dubois D, Pereira B, et al. 2014. Colonization of the human gut by E. coli and colorectal cancer risk. *Clin Cancer Res*.20(4):859-67.

8. Zhao L-Y, Mei J-X, Yu G, Lei L, Zhang W-H, Liu K, et al. 2023. Role of the gut microbiota in anticancer therapy: from molecular mechanisms to clinical applications. *Signal Transduction and Targeted Therapy*.8(1):201.

9. Sprouffske K, Aguilar-Rodríguez J, Sniegowski P, Wagner A. 2018. High mutation rates limit evolutionary adaptation in Escherichia coli. *PLoS Genet*.14(4):e1007324.

10. Wang Y, Fu K. 2023. Genotoxins: The Mechanistic Links between Escherichia coli and Colorectal Cancer. *Cancers (Basel)*.15(4).

11. Johnson JR, Johnston B, Kuskowski MA, Nougayrede JP, Oswald E. 2008. Molecular epidemiology and phylogenetic distribution of the Escherichia coli pks genomic island. *J Clin Microbiol*.46(12):3906-11.

12. Baker EP, Sayegh R, Kohler KM, Borman W, Goodfellow CK, Brush ER, Barber MF. 2022. Evolution of host-microbe cell adherence by receptor domain shuffling. *eLife*.11:e73330.

13. Nouri R, Hasani A, Shirazi KM, Sefiadn FY, Mazraeh FN, Sattarpour S, Rezaee MA. 2024. Colonization of the gut mucosa of colorectal cancer patients by pathogenic mucosa-associated Escherichia coli strains. *Diagnostic Microbiology and Infectious Disease*.109(2):116229.

14. Paixão AC, Ferreira AC, Fontes M, Themudo P, Albuquerque T, Soares MC, et al. 2016. Detection of virulence-associated genes in pathogenic and commensal avian Escherichia coli isolates. *Poultry Science*.95(7):1646-52.

15. Nouri R, Hasani A, Masnadi Shirazi K, Alivand MR, Sepehri B, Sotoudeh S, et al. 2021. Mucosa-Associated Escherichia coli in Colorectal Cancer Patients and Control Subjects: Variations in the Prevalence and Attributing Features. *Can J Infect Dis Med Microbiol*.2021:2131787.

16. Pokharel P, Dhakal S, Dozois CM. The Diversity of Escherichia coli Pathotypes and Vaccination Strategies against This Versatile Bacterial Pathogen. Microorganisms [Internet]. 2023; 11(2).

17. Raisch J, Buc E, Bonnet M, Sauvanet P, Vazeille E, de Vallée A, et al. 2014. Colon cancer-associated B2 Escherichia coli colonize gut mucosa and promote cell proliferation. *World J Gastroenterol*.20(21):6560-72.

18. Pokharel P, Dhakal S, Dozois CM. 2023. The Diversity of Escherichia coli Pathotypes and Vaccination Strategies against This Versatile Bacterial Pathogen. *Microorganisms*.11(2).

[DOI: 10.61186/ijbc.16.3.63]