ISSN 0891-4168, Molecular Genetics, Microbiology and Virology, 2016, Vol. 31, No. 1, pp. 15–20. © Allerton Press, Inc., 2016.
Original Russian Text © N.T. Nguyen, R.R. Vafin, I.V. Rzhanova, A.I. Kolpakov, I.G. Gataullin, S.V. Tyulkin, M.N. Sinyagina, T.V. Grigoryeva, O.N. Ilinskaya, 2016, published in Molekulyarnaya Genetika, Mikrobiologiya i Virusologiya, 2016, No. 1, pp. 13–18.

EXPERIMENTAL WORKS

Molecular Genetic Analysis of Microorganisms with Intraepithelial Invasion Isolated from Patients with Colorectal Cancer

N. T. Nguyen[a], R. R. Vafin[a, b], I. V. Rzhanova[a], A. I. Kolpakov[a], I. G. Gataullin[c, d], S. V. Tyulkin[e], M. N. Sinyagina[a], T. V. Grigoryeva[a], and O. N. Ilinskaya[a]

[a] *Kazan Federal University, Kazan, Tatarstan, 420008 Russia*
[b] *Tatar Research Institute of Agriculture, Russian Academy of Sciences, Kazan, Tatarstan, 420059 Russia*
[c] *Tatarstan Regional Clinical Cancer Center, Kazan, Tatarstan, 420002 Russia*
[d] *Kazan State Medical Academy, Kazan, Tatarstan, 420012 Russia*
[e] *Tatar Interregional Veterinary Laboratory, Kazan, Tatarstan, 420087 Russia*
e-mail: Olga.Ilinskaya@kpfu.ru

Received April 16, 2015

Abstract—The facultative aerobic bacteria isolated from the mucosa of rectum in the zone of malignant tumor and neighboring normal mucosa in patients with colorectal cancer were studied using molecular genetic methods. The species attribution of bacteria was implemented using the cultural morphological analysis and sequencing of the 16S rRNA locus. The microorganisms with intraepithelial invasion to rectal mucosa were identified as the representatives of adherent-invasive (AIEC) subgroup of *Escherichia coli* and species *Klebsiella pneumonia*. The molecular analysis by genetic determinants controlling adhesive, hemolytic, and toxigenic activity revealed that some bacterial isolates were able to produce toxins with potential cancerogenic activity (e.g., colibactin and cytotoxic necrotizing factor 1). Certain bacterial species isolated from malignant and normal rectal epithelium of the same patient demonstrated no difference between analyzed factors of toxigenicity.

Keywords: colorectal cancer, *Escherichia coli*, *Klebsiella pneumonia*, colibactin, cytotoxic necrotizing factor 1, identification, PCR, sequencing

DOI: 10.3103/S0891416816010079

INTRODUCTION

Microflora of human intestine contributes significantly to the process of the onset and development of colorectal cancer. The mechanisms of carcinogenesis, associated with the microbial community, include the genotoxic effect of a number of bacterial toxins such as colibactin, cytolethal distending toxin (CDT), cycle inhibiting factor (CIF), and cytotoxic necrotizing factor 1 (CNF1) [1, 2]. Different types of microorganisms are capable of producing the aforementioned toxins. Thus, colibactin leading to oncogenic mutations due to double-strand DNA breaks is produced by the representatives of species *Klebsiella pneumoniae*, *Enterobacter aerogenes*, and *Citrobacter koseri*, belonging to the family Enterobacteriaceae, in addition to several strains of *Escherichia coli* [3].

Significant arrayed data demonstrate the ability of adherent-invasive *E. coli* strains to colonize the colon mucosa of patients with colorectal cancer [2, 4]. Moreover, it has been shown that the number of microorganisms of opportunistic pathogenic microflora is increased in 48% of patients with colorectal cancer, while *K. pneumoniae* (13.5%), *Enterobacter* spp. (7.6%), and *Citrobacter* spp. (3.8%) [5], i.e., bacteria that are able to synthesize genotoxic metabolites, were indentified among the intestinal microbial community members [3]. In this regard, the idea that the molecular basis of epidemiological link between colorectal cancer and *K. pneumoniae*, which actively colonizes the intestinal mucosa [6] and produces colibactin [7], is determined by the carcinogenic activity of colibactin [3], appears quite reasonable.

In this study, molecular genetic analysis of the facultative aerobic bacteria isolated from the biopsies of malignant and normal rectal epithelium obtained during surgical procedures in patients with colorectal cancer was carried out. The aim of the study was to identify and compare the genetic determinants that control adhesive, hemolytic, and toxigenic activity in dominant isolates of *E. coli* associated with normal and malignantly transformed colon epithelium.

MATERIALS AND METHODS

Six samples of microbial cultures isolated on meat infusion agar from malignant tumor of rectal mucosa

Table 1. List of oligonucleotide primers used in the study

No.	Gene locus	Oligonucleotide primer, its sequence and length (b.)	PCR product (bp)	Reference
1	16S rRNA	ITS-F: 5'-GATTAGATACCCTGGTAG-3' (18 b.) ITS-R: 5'-AGTCACTTAACCATACAACCC-3' (21 b.)	1215	[8]
2	*papC*	papC-F: 5'-CGTTCGCCGGGTATCGTTTCTCAG-3' (24 b.) papC-R: 5'-CCCGTTCCCCAGCGATTTGTCAC-3' (23 b.)	724	[9]
3	*papH*	papH-F: 5'-TTAAAGATAATCGGGTCAT-3' (19 b.) papH-R: 5'-GGAATCAGAGAAAAGGTT-3' (18 b.)	858	[9]
4	*afa*	afa-f: 5'-GCTGGGCAGCAAACTGATAACTCTC-3' (25 b.) afa-r: 5'-CATCAAGCTGTTTGTTCGTCCGCCG-3' (25 b.)	750	[10]
5	*eae*	Eae-f: 5'-TCAATGCAGTTCCGTTATCAGTT-3' (23 b.) Eae-r: 5'-GTAAAGTCCGTTACCCCAACCTG-3' (24 b.)	482	[11]
6	*bfpA*	bfpA_114F: 5'-GTCTGCGTCTGATTCCAATA-3 (20 b.) bfpA_521R: 5'-TCAGCAGGAGTAATAGC-3' (17 b.)	408–114	[12]
7	*hlyA*	hlyA-F: 5'-GCATCATCAAGCGTACGTTCC-3' (21 b.) hlyA-R: 5'-AATGAGCCAAGCTGGTTAAGCT-3' (22 b.)	534	[13]
8	*lthB*	LT-f: 5'-ACGGCGTTACTATCCTCTC-3' (19 b.) LT-r: 5'-TGGTCTCGGTCAGATATGTG-3' (20 b.)	273	[11]
9	*sta*	STa-1: 5'-GCTAATGTTGGCAATTTTTATTTCTGTA-3' (28 b.) STa-2: 5'-AGGATTACAACAAAGTTCACAGCAGTAA-3' (28 b.)	190	[14]
10	*stx*	VTcom-f: 5'-GAGCGAAATAATTTATATGTG-3' (21 b.) VTcom-r: 5'-TGATGATGGCAATTCAGTAT-3' (20 b.) (294)	518	[11]
11	*pks*	pks-f: 5'-ATTCGATAGCGTCACCCAAC-3' (20 b.) pks-r: 5'-TAAGCGTCTGGAATGCAGTG-3' (20 b.)	2119	[2]
12	*cnfl*	CNF-1s: 5'-GGGGGAAGTACAGAAGAATTA-3' (21 b.) CNF-1as: 5'-TTGCCGTCCACTCTCACCAGT-3' (21 b.)	1112	[2]
13	*cdtB-I*	CDT-I-s: 5'-CAATAGTCGCCCACAGGA-3' (18 b.) CDT-I-as: 5'-ATAATCAAGAACACCACCAC-3' (20 b.)	411	[2]
14	*cdtB-IV*	CDT-IV-s: 5'-CCTGATGGTTCAGGAGGCTGGTTC-3' (24 b.) CDT-IV-as: 5'-TTGCTCCAGAATCTATACCT-3' (20 b.)	350	[2]
15	*cif*	cif-int-s: 5'-AACAGATGGCAACAGACTGG-3' (20 b.) cif-int-as: 5'-AGTCAATGCTTTATGCGTCAT-3' (21 b.)	383	[2]

(samples Tru26o, Shi1o, Ef15o) and normal rectal mucosa (samples Tru26n, Shi1n, Ef15n) from three patients with colorectal cancer were investigated in pairs. Biopsy sampling was performed in accordance with the resolution of the Ethics Committee of Kazan State Medical Academy (Minute no. 4 dated May 7, 2009). The biopsy samples were incubated for 1 h in 1 mL of sterile saline at 37°C; then, 100 μL of the solution was inoculated on meat infusion agar and grown until colonies formed. Genomic DNA was isolated from the cultures of microorganisms using a DNA-sorb-B kit (Central Research Institute of Epidemiology, Federal Service for Veterinary and Phytosanitary Surveillance). The samples were stored at –20°C. Polymerase chain reaction (PCR) was performed on a MJ Mini Gradient Thermal Cycler thermocycler (Bio-Rad, United States). The detection of the results of PCR analysis was carried out by horizontal electrophoresis in 2.5% agarose gel in TBE buffer (pH 8.0) containing ethidium bromide, followed by visualization of the results using an ultraviolet transilluminator ($\lambda = 310$ nm). The length of DNA fragments was evaluated in comparison with standard DNA markers.

Identification of genetic determinants controlling adhesive (*papC, papH*, *papH*, *afa*, *eae*, *bfpA*), hemolytic (*hlyA*), and toxigenic (*lthB*, *STa*, *Stx*, *pks*, *cnfI*, *cdtB*, *cif*) activity was performed using primers selected according to the literature mostly for *E. coli* gene diagnostics and represented in Table 1.

Reagents for molecular biological studies produced by DNA synthesis (Russia), SibEnzyme (Russia), and Applied Biosystems (United States) were used.

Sequencing of loci 16S rRNA, *pks* and *cnf1* amplification products was performed on a genetic analyzer

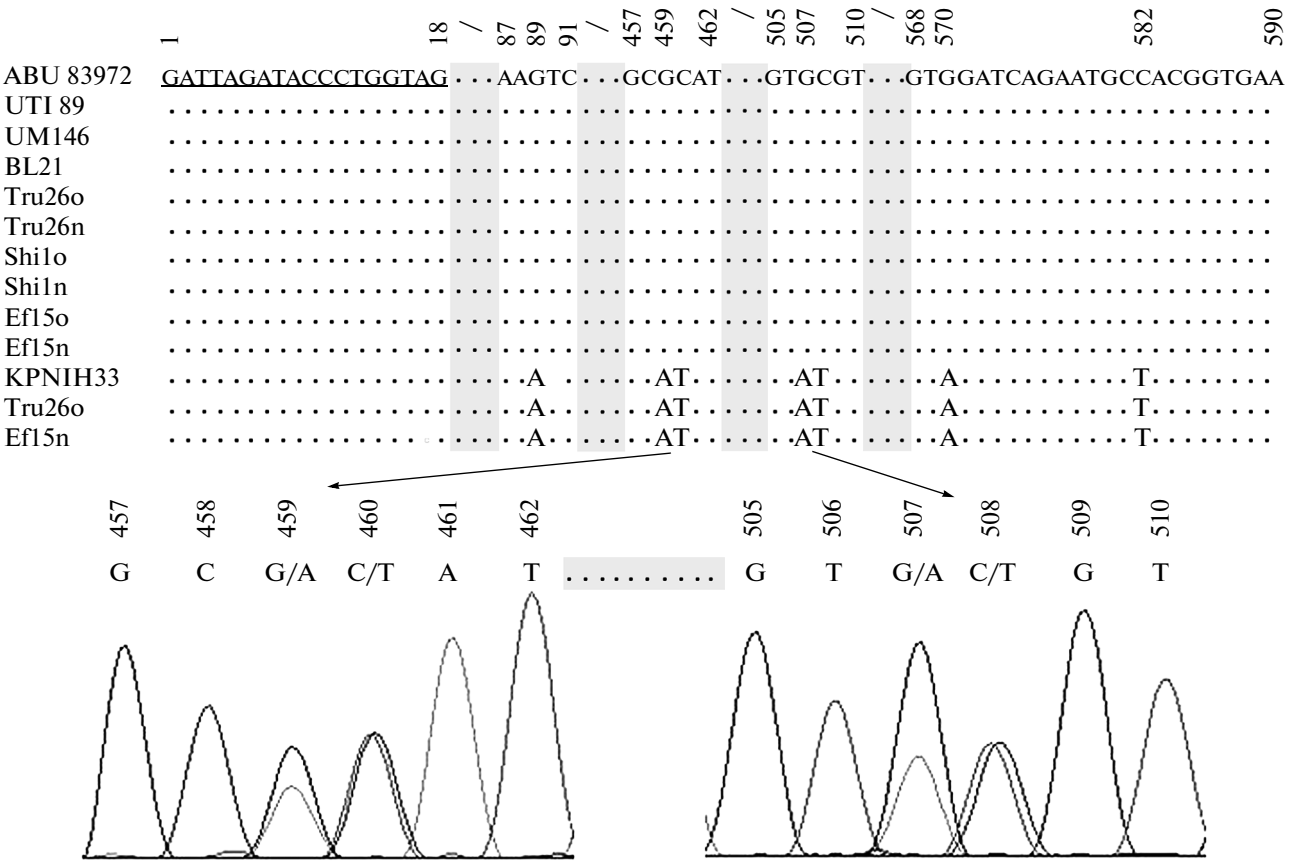

Fig. 1. SNP positions in the analyzed 16S rRNA loci of *E. coli* and *K. pneumoniae*.

ABI PRISM 3100 (Applied Biosystems, United States). The sequences of the loci 16S rRNA, *pks* and *cnf1* were aligned according to the nucleotide sequences of a number of reference microorganism strains published in GenBank using the programs BLAST and CLUSTAL W (v. 1.83).

RESULTS AND DISCUSSION

Examination of biopsy bacterization revealed that the number of facultative anaerobic heterotrophs closely associated with the tissue of rectal mucosa is small, and it corresponds to a value of up to 80 colony-forming units per medium-sized biopsy (about 1–8 mm^3). Based on morphological and cultural characteristics of isolated bacterial cultures, three pairs of typical isolates from three patients were identified as *E. coli*. Sequencing of the 16S rRNA locus of six studied isolates confirmed the species attribution of Tru26n, Shi1o, Shi1n, and Ef15o to *E. coli*. Comparison of identificational SNP positions in the analyzed loci of 16S rRNA in the studied samples of microbial cultures and reference strains of *E. coli* (ABU 83972, UT189, UM146, BL21) and *K. pneumonia* (KPNIH33) revealed the presence of double fluorescence peaks at positions 459–460 and 507–508 for the Tru26o and Ef15n samples (Fig. 1), and it can be concluded that these samples represent a mixed culture of the microorganisms *E. coli* and *K. pneumoniae* growing in close association.

Numerical values of SNP positions in the analyzed loci of 16S rRNA were specified by the sequence of the

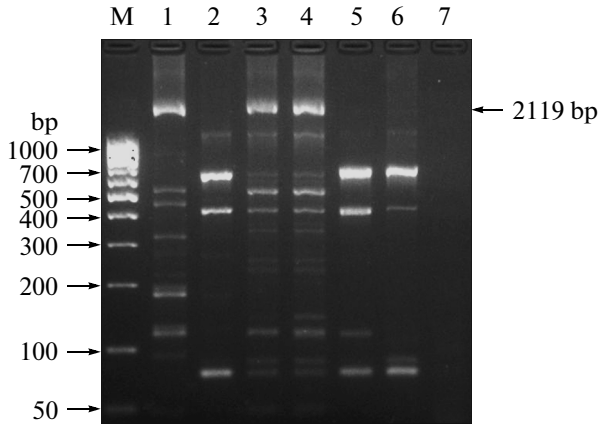

Fig. 2. Electrophoregram of PCR result with primers pks-f + pks-r (amplification of the *pks* gene locus of *E. coli* and *K. pneumoniae*). Designations: (1) Tru26o, (2) Tru26n, (3) Shi1o, (4) Shi1n, (5) Ef15o, (6) Ef15n+, and (7) negative control.

```
                   1                   20 /  1039 1041 1043 /  1738 1740 1742 /  1852 1854 1856 /  2100                  2119
KPNIH33            ATTCGATAGCGTCACCCAAC...CGTGC...GCATC...AGACC...CACTGCATTCCAGACGCTTA
ATCC BAA–895       ......................................................................
EA1509E            ......................................................................
Tru26o             ......................................................................
ABU 83972          ..........................G.......G.......................G.........
UTI 89             ..........................G.......G.......................G.........
UM146              ..........................G.......G.......................G.........
Shi1o              ......................A.....G.......G.......................G.........
Shi1n              ......................A.....G.......G.......................G.........
```

Fig. 3. SNP positions in the analyzed *pks* gene loci of *E. coli*, *K. pneumoniae*, *C. koseri*, and *E. aerogenes*.

primer ITS-F: 5'-GATTAGATACCCTGGTAG-3' (positions 1–18), which was also used as a sequence oligonucleotide.

Two fluorescence peaks superimposed on each other in positions 459–460 and 507–508 for the samples Tru26o and Ef15n are a clear demonstration of the fact that each of them is indeed a mixed culture (see Fig. 1).

When analyzing the marker genes, products of which may contribute to carcinogenesis, it was found that the samples Tru26o, Shi1o, and Shi1n gave a positive amplification signal for the locus of the *pks* gene encoding colibactin (Fig. 2).

The sample of the mixed culture Ef15n also initially generated specific amplicon of the *pks* gene locus. However, the pure culture *E. coli* (Ef15n*) isolated from the mixed culture lost this property; its DNA no longer led to amplification of this PCR product (see Fig. 2, track 6), which indicates the attribution of the *pks* gene to the genome of *K. pneumoniae*, the second associate of this mixed culture.

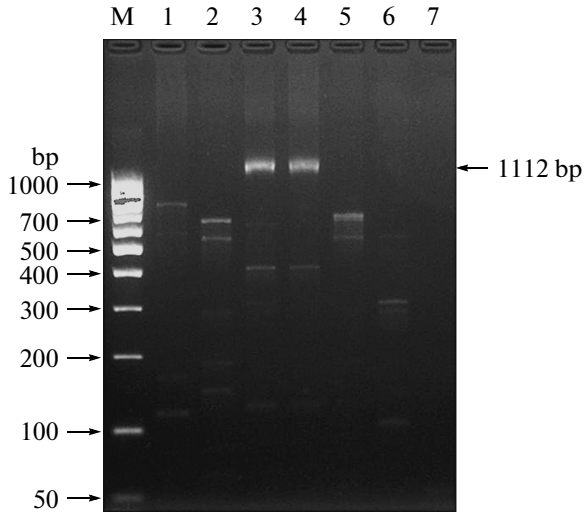

Fig. 4. Electrophoregram of PCR result with primers CNF-1s+CNF-1as (amplification of the *cnf1* gene locus of *E. coli*). Designations: (1) Tru26o, (2) Tru26n, (3) Shi1o, (4) Shi1n, (5) Ef15o, (6) Ef15n, and (7) negative control.

Sequencing of the specific PCR product (2119 bp) and subsequent alignment of the corresponding DNA sequences revealed that the samples Shi1o and Shi1n had identical structure of the analyzed *pks* gene locus, which differed in just one synonymous nucleotide replacement (SNP position 1041) compared with the reference *E. coli* strains ABU 83972, UTI89, and UM146 (Fig. 3).

The nucleotide sequence of the *pks* gene locus of the mixed culture sample Tru26o belonging to the genome of *K. pneumoniae* was similar to the corresponding sequences of the reference strains of the species *K. pneumoniae* (KPNIH33 strain), *C. koseri* (ATCC BAA-895 strain), and *E. aerogenes* (EA1509E strain), which are also characterized by the presence of the *pks* gene (see Fig. 3). The isolate *E. coli* (Tru26o*) isolated from the mixed culture did not possess the ability to produce colibactin like Ef15n*.

The *E. coli* samples Shi1o and Shi1n had a positive amplification signal for the *cnf1* (cytotoxic necrotizing factor 1) gene locus, which was confirmed by sequencing of the specific PCR product with a length of 1112 bp (Fig. 4).

These samples had the structure of the analyzed *cnf1* gene locus identical to each other, as well as to the reference *E. coli* strain ABU 83972, which was determined as a result of the alignment of corresponding sequenced DNA sequences (Fig. 5).

Thus, among the six isolates of *E. coli* and two isolates of *K. pneumoniae* (obtained during the separation of the mixed cultures Tru26o and Ef15n), both isolates of *K. pneumoniae* and two isolates of *E. coli* are capable of colibactin synthesizing, while these two *E. coli* isolates can also synthesize cytotoxic necrotizing factor 1.

According to the results of PCR analysis, the carriers of none of the following genes were found in any of the studied bacteria: *papC*, *papH*, *afa*, *eae*, *bfpA*, *hlyA*, *lthB*, *sta*, *stx*, *cdtBI*, *cdtBIV*, and *cif*.

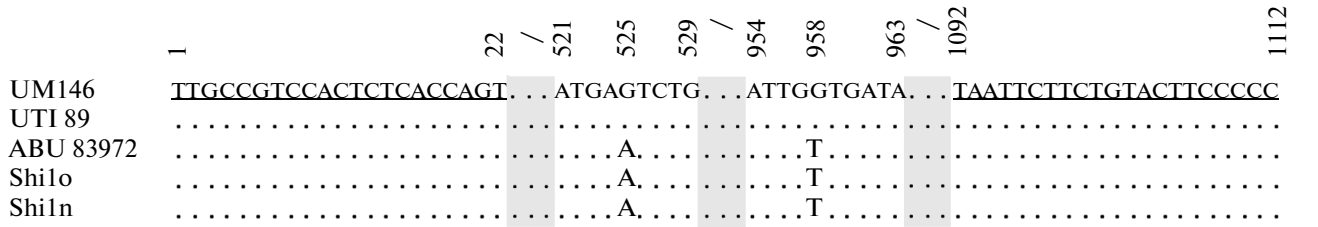

Fig. 5. SNP positions in the analyzed *cnf1* gene loci of *E. coli*.

A summary of the molecular genetic analysis of the bacteria isolated from patients with colorectal cancer is represented in Table 2.

The absence of the amplification signal for a number of analyzed loci excluded the attribution of the studied bacteria to enteropathogenic (EPEC), enterotoxigenic (ETEC), enterohemorrhagic (EHEC), shigalike toxin-producing (STEC), and uropathogenic (UPEC) subgroups of *E. coli*.

It should be noted that the bacteria from the samples of malignant and normal epithelium of colonic mucosa in the same patient do not differ in the presence of toxigenicity factor genes, which were analyzed, although this difference was documented between the cultures isolated from different patients (see Table 2). Based on these results the following conclusions were drawn.

1. The microorganisms with pronounced intraepithelial invasion to rectal mucosa isolated from patients with colorectal cancer were identified as representatives of adherent-invasive (AIEC) subgroup of *E. coli* and the species *K. pneumoniae*.

2. The presence of isolates producing carcinogenic toxins such as colibactin and cytotoxic necrotizing factor 1, which are pathogenic oncomarkers, in the rectal mucosa of patients with colorectal cancer was determined by molecular analysis of the isolated microbial cultures.

3. The fact of identification of colibactin-producing isolates, whether *Escherichia coli* or *Klebsiella pneumoniae*, in all the patients is consistent with an epidemiological link between the identified microorganisms and colorectal cancer onset, which should be considered in further improvement of the schemes of diagnostic and monitoring studies with respect to this cancer.

ACKNOWLEDGMENTS

This work was performed within the Russian Government Program of Competitive Growth of Kazan Federal University using equipment of the Interdisciplinary Center of Kazan Federal University (ID RFMEFI59414X0003).

Table 2. Summary of the molecular genetic analysis of microbial isolates isolated from patients with colorectal cancer

Species	Isolate	Analyzed gene loci														
		16S rRNA	*papC*	*papH*	*afa*	*eae*	*bfpA*	*hlyA*	*lthB*	*sta*	*stx*	*cnfI*	*pks*	*cdtB I*	*cdtB IV*	*cif*
K. pneumoniae	Tru260	+	–	–	–	–	–	–	–	–	–	–	+	–	–	–
E. coli	Tru26o*	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–
E. coli	Tru26n	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–
E. coli	Shi1o	+	–	–	–	–	–	–	–	–	–	+	+	–	–	–
E. coli	Shi1n	+	–	–	–	–	–	–	–	–	–	+	+	–	–	–
E. coli	Efl5o	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–
K. pneumoniae	Efl5n	+	–	–	–	–	–	–	–	–	–	–	+	–	–	–
E. coli	Ef15n*	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Designations: "+"—the presence of analyzed gene locus; "–"—the absence of analyzed gene locus.

REFERENCES

1. Kutikhin, A.G., Brusina, E.B., Popova, Yu.A., Yuzhalin, A.E., Zhivotovskiy, A.S., and Briko, N.I., Correlation between the colonic microflora and the genesis of colorectal cancer, *Epidemiol. Infekts. Bolezn. Aktual. Vopr.,* 2012, no. 4, pp. 44–52.
2. Raisch, J., Buc, E., Bonnet, M., Sauvanet, P., Vazeille, E., de Vallee, A., et al., Colon cancer-associated B2 *Escherichia coli* colonize gut mucosa and promote cell proliferation, *World J. Gastroenterol.,* 2014, vol. 20, no. 21, pp. 6560–6572.
3. Putze, J., Hennequin, C., Nougayrede, J.P., Zhang, W., Homburg, S., Karch, H., et al., Genetic structure and distribution of the colibactin genomic island among members of the family Enterobacteriaceae, *Infect. Immun.,* 2009, vol. 77, no. 11, pp. 4696–4703.
4. Swidsinski, A., Khilkin, M., Kerjaschki, D., Schreiber, S., Ortner, M., Weber, J., et al., Association between intraepithelial *Escherichia coli* and colorectal cancer, *Gastroenterology,* 1998, vol. 115, no. 2, pp. 281–286.
5. Starostina, M.A., Afanas'eva, Z.A., Gubaeva, M.S., Ibragimova, N.R., and Sakmarova, L.I., Colon biocoenosis in patients with colorectal cancer, *Prakt. Med.,* 2012, vol. 66, no. 6, pp. 97–99.
6. Hennequin, C. and Forestier, C., OxyR, a LysR-type regulator involved in Klebsiella pneumonia mucosal and abiotic colonization, *Infect. Immun.,* 2009, vol. 77, no. 12, pp. 5449–5457.
7. Lai, Y.C., Lin, A.C., Chiang, M.K., Dai, Y.H., Hsu, C.C., Lu, M.C., et al., Genotoxic *Klebsiella pneumonia* in Taiwan, *PLoS One,* 2014, vol. 9, no. 5, p. e96292. http://dx.plos.org/10.1371/journal.pone.0096292
8. Magray, M.S., Kumar, A., Rawat, A.K., and Srivastava, S., Identification of *Escherichia coli* through analysis of 16S rRNA and 16S-23S rRNA internal transcribed spacer region sequences, *Bioinformation,* 2011, vol. 6, no. 10, pp. 370–371.
9. Krasnoperova, Yu.Yu. and Aseinova, R.V., Some molecular and genetic mechanisms of mutual adaptation of *Escherichia coli* strains before and after coculture with the *Blastocystis hominis*, *Vestn. Nov. Med. Tekhnol.,* 2009, vol. 16, no. 1, pp. 12–14.
10. Liu, P., Soupir, M.L., Zwonitzer, M., Huss, B., and Jarboe, L.R., Association of antibiotic resistance in agricultural *Escherichia coli* isolates with attachment to quartz, *Appl. Environ. Microbiol.,* 2011, vol. 77, no. 19, pp. 6945–6953.
11. Tobias, J. and Vutukuru, S.R., Simple and rapid multiplex PCR for identification of the main human diarrheagenic *Escherichia coli*, *Microbiol. Res.,* 2012, vol. 167, no. 9, pp. 564–570.
12. Lacher, D.W., Steinsland, H., Blank, T.E., Donnenberg, M.S., and Whittam, T.S., Molecular evolution of typical enteropathogenic *Escherichia coli*: clonal analysis by multilocus sequence typing and virulence gene allelic profiling, *J. Bacteriol.,* 2007, vol. 189, no. 2, pp. 342–350.
13. Hegde, A., Ballal, M., and Shenoy, S., Detection of diarrheagenic *Escherichia coli* by multiplex PCR, *Indian J. Med. Microbiol.,* 2012, vol. 30, no. 3, pp. 279–284.
14. Kim, Y.J., Kim, J.H., Hur, J., and Lee, J.H., Isolation of *Escherichia coli* from piglets in South Korea with diarrhea and characteristics of the virulence genes, *Can. J. Vet. Res.,* 2010, vol. 74, no. 1, pp. 59–64.

Translated by D. Novikova