ISSN: 1674-0440

ORIGINAL ARTICLE

Phylotyping and Genotyping of *Escherichia Coli* Isolates Obtained From Patients with Colorectal Cancer

Irina V. Rzhanova[1], Ramil R. Vafin[2], Tkhi Nga Nguyen[3], Alexey I. Kolpakov[1], Olga N. Ilinskaya[1]

[1]*Kazan Federal University*
[2]*V.M. Gorbatov Federal Research Center for Food Systems of Russian Academy of Sciences*
[3]*Hanoi Medical University*

Abstract

Objective of the research was to phylotype the *E. coli* isolates from the malignant and adjacent normal epithelium of the rectal mucosa of patients diagnosed with colorectal cancer, identifying and discriminating phylogroups by a set of genes, and also genotyping them by the *fimH*-gene locus by identifying *fimH* types. Samples of bacterial cultures isolated from paired biopsy specimens of patients diagnosed with colorectal cancer were subjected to molecular genetic research for species identification, as well as *E. coli* phylo- and genotyping. By sequencing the 16S *rRNA* locus, bacterial cultures are identified both as isolates of monocultures of *E. coli* and as mixed cultures of *E. coli* and *K. pneumonia*. The procedure for the phylotyping of *E. coli* isolates by identifying and discriminating phylogroups by a set of genes established the affiliation of samples of bacterial cultures with phylogroups A and B2. The genotyping procedure for *E. coli* isolates at the *fimH* gene locus by identifying the *fimH* types characterized the studied samples by signs of the third (f-3), fourth (f-4), and seventh (f-7) types, respectively. The strategy for identifying *fimH*-types during genotyping of *E. coli* at the *fimH*-gene locus has been improved, based on the analysis of SNP sequenced DNA sequences, and includes 52 *fimH*-types, three of which are justified in this paper.

Keywords: colorectal cancer, *Escherichia coli*, phylotyping, genotyping, *fimH*, PCR, sequencing, SNP

1. INTRODUCTION

Microorganisms capable of colonizing the mucous membrane of the large intestine may be etiological agents of the onset and development of colorectal cancer, with the mechanism of carcinogenesis, also caused by the genotoxic effect of a number of cyclomodulin toxins produced by them, in particular, olibactin, resulting in oncogenic mutations due to the double helix breaks in the DNA [1, 2, 3].

An epidemiological link has been identified between colorectal cancer and certain members of the *Enterobacteriaceae* family, including *Escherichia coli* and *Klebsiella pneumoniae*, actively colonizing the intestinal mucosa and producing colibactin [4, 5, 6].

It was established that the main part of identified cyclomodulin-positive *E. coli* isolated from biopsies of malignant epithelium of the rectal mucosa of patients with a diagnosis of colorectal cancer belongs to the phylogroup B2, therefore phylotyping of *E. coli* isolates can be a diagnostically significant procedure for cancer screening [1, 2, 7].

The proposed strategy of identification and discrimination of phylogroups by a set of genes allows for the effective identification of 7 known phylogroups [8]. This procedure of phylotyping *E. coli* in combination with their *fimH* locus genotyping by identifying *fimH* types is important in the molecular epidemiology of infections, including those associated with oncology. However, based on the analysis of single nucleotide polymorphism of sequenced DNA sequences, the strategy of genotyping *E. coli* for the *fimH* gene locus included from 46 to 49 *fimH* types [9, 10].

Correspondence Author: Irina V. Rzhanova
Email address: *rzh.irina@mail.ru*
Received: 21-06-2018
Revised: 15-08-2018
Accepted:21-09-2018

FimH gene is an informative target. It is known that adhesion of *E. coli* to epithelial cells is mediated by fibrobria carrying a specific protein - adhesin - *FimH*, regulating tissue tropism, up to and beyond the limits of the primary ecological niche, due to point mutations leading to an increase in the pathogenic potential of the microorganism [11].

Objective of the research was to phylotype the *E. coli* isolates from the malignant and adjacent normal epithelium of the rectal mucosa of patients diagnosed with colorectal cancer, identifying and discriminating phylogroups by a set of genes, and also genotyping them by the *fimH*-gene locus by identifying *fimH* types.

2. Methods

A molecular genetic study was performed on 6 samples of bacterial cultures isolated on meat-peptone agar (MPA) from paired biopsy specimens taken from three patients diagnosed with colorectal cancer. Samples of bacterial cultures Tru26o, Shi1o and Ef15o were isolated from the malignant epithelium of the rectal mucosa, and samples Tru26n, Shi1n and Ef15n from the adjacent normal epithelium, respectively. Biopsy samples selected in accordance with the Resolution of the Ethical Committee of the Kazan State Medical Academy (protocol No. 4 dated May 7, 2009) were incubated for 1 hour at 37°C in 1 ml of sterile saline, followed by plating 100 µl of solution on MPA and growing cultures of bacteria until the colonization. Extraction of genomic DNA from samples of bacterial cultures was carried out with the DNA-Sorb B kit (Central Research Institute of Epidemiology, Russia). Specific identification of isolated bacterial cultures, as well as phylo- and genotyping of *E. coli*, was performed using synthesized oligonucleotide primers (DNA synthesis, Russia), presented in Table 1. Polymerase chain reaction (PCR) formulation was performed on an MJ Mini Gradient Thermal Cycler amplifier (Bio-Rad, USA). Electrophoretic detection of PCR amplification products was performed with an EF-genotip 200 reagent kit (Central Research Institute of Epidemiology, Russia) in a 2.5% agarose gel in TBE buffer containing ethidium bromide, followed by visualization of the amplicons in a UV transilluminator (λ=310 nm) whose sizes were compared with standard DNA markers (SibEnzim, Russia). The sequencing of the amplicons of the *E. coli 16S rRNA* and *fimH* loci was performed on an ABI PRISM 3100 genetic analyzer (Applied Biosystems, USA), followed by their alignment in BLAST and CLUSTAL W (v. 1.83) with the corresponding microbial nucleotide sequences previously published in GenBank NCBI

Table 1. A list of oligonucleotide primers used in the work

No.	Gene locus	Oligonucleotide primers, their sequence and length (n.)	PCR product (bp)	Reference
Primers for the specific identification of bacteria by direct amplicon sequencing				
1	*16S rRNA*	ITS-F: 5′-GATTAGATACCCTGGTAG-3′ (18 n.) ITS-R: 5′-AGTCACTTAACCATACAACCC-3′ (21 n.)	1215	[12, 13]
Primers for *E. coli* phylotyping with quadruplex PCR				
2	*arpA*	AceK.f: 5′-AACGCTATTCGCCAGCTTGC-3′ (20 n.) ArpA1.r: 5′-TCTCCCCATACCGTACGCTA-3′ (20 n.)	400	[8, 14]
3	*chuA*	chuA.1b: 5′-ATGGTACCGGACGAACCAAC-3′ (20 n.) chuA.2: 5′-TGCCGCCAGTACCAAAGACA-3′ (20 n.)	288	[8, 15]
4	*yjaA*	yjaA.1b: 5′-CAAACGTGAAGTGTCAGGAG-3′ (20 n.) yjaA.2b: 5′-AATGCGTTCCTCAACCTGTG-3′ (20 n.)	211	[8]
5	TspE4.C2	TspE4C2.1b: 5′-CACTATTCGTAAGGTCATCC-3′ (20 n.) TspE4C2.2b: 5′-AGTTTATCGCTGCGGGTCGC-3′ (20 n.)	152	[8]
Primers for *E. coli* phylogroups E and C discrimination				
6	*trpA*	C-specific primers: trpAgpC.1: 5′-AGTTTTATGCCCAGTGCGAG-3′ (20 n.) trpAgpC.2: 5′-TCTGCGCCGGTCACGCCC-3′ (18 n.)	219	[8, 16]
7	*arpA*	E-specific primers: ArpAgpE.f: 5′-GATTCCATCTTGTCAAAATATGCC-3′ (24 n.) ArpAgpE.r: 5′-GAAAAGAAAAAGAATTCCCAAGAG-3′ (24 n.)	301	[8, 16]
8	*trpA*	Internal control: trpBA.f: 5′-CGGCGATAAAGACATCTTCAC-3′ (21 n.) trpBA.r: 5′-GCAACGCGGCCTGGCGGAAG-3′ (20 n.)	489	[8, 17]

Primers for *E. coli* genotyping by *fimH* gene sequencing				
9	*fimH*	FimH-f: 5/-CGAGTTATTACCCTGTTTGCTG-3/ (22 н.) FimH-r: 5/-ACGCCAATAATCGATTGCAC-3/ (20 н.)	878	[9, 10]

3. Results and Discussions

Previously, based on the cultural morphological characteristics and sequencing results of the *16S rRNA* locus, the samples of bacterial cultures Tru26n, Shi1o, Shi1n, Ef15o were identified as isolates of monocultures of *Escherichia coli*; Tru26o and Ef15n were an association of *Escherichia coli* and *Klebsiella pneumonia* microorganisms [13].

In the present work, the phylotyping of isolates with quadruplex PCR was guided by the strategy of identification and discrimination of phylogroups by the set of genes shown in Table 2.

Table 2. The strategy of identification and discrimination of phylogroups for phylotyping *E. coli* by the set of genes

arpA **(400 bp)**	***chuA*** **(288 bp)**	***yjaA*** **(211 bp)**	**TspE4.C (152 bp)**	**Phylogroup**	**Next step**
+	-	-	-	**A**	
+	-	+	-	**A or C**	**Discrimination of phylogroups with C-specific primers**
+	-	-	+	**B1**	
-	+	+	-	**B2**	
-	+	-	+		
-	+	+	+		
+	+	-	-	**D or E**	**Discrimination of phylogroups with E-specific primers**
+	+	-	+		
+	+	+	-	**E**	
-	+	-	-	**F**	

As a result of quadruplex PCR for *E. coli* phylotyping, isolates of Shi1o and Shi1n samples were identified as representatives of phylogroup B2, and samples of Tru26o, Tru26n, Ef15o, and Ef15n generated a spectrum of phylogroup-specific fragments that are simultaneously characteristic of phylogroups A and C (Fig. 1).

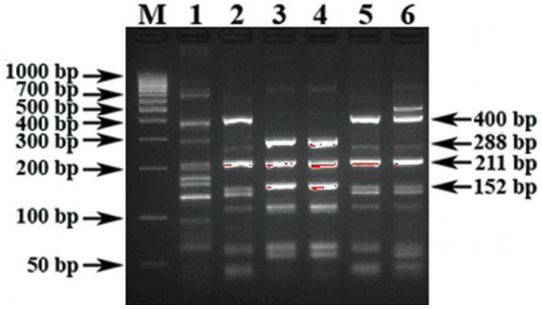

Fig. 1. Electrophoregram of the resulting quadruplex PCR for phylotyping *E. coli*
Designations: 1) Tru26o; 2) Tru26n; 3) Shi1o; 4) Shi1n; 5) Ef15o; 6) Ef15n.

Four pairs of primers used in the reaction also initiated the amplification of minor non-specific fragments that did not significantly affect the correct interpretation of the obtained PCR result (Fig. 1).

Moreover, if the *E. coli* isolates of the Shi1o and Shi1n samples were identified by amplification characteristic of the phylogroup B2 (288/211/152 bp) directly in the formulation of quadruplex PCR (Table 1-2, Figure 1), the other samples The phylogroup A was established as the next step - PCR to discriminate phylogroups with C-specific primers, which did not lead to amplification of the 219 bp-long *trpA* gene locus (Table 1-2).

The fimH locus genotyping of *E. coli* isolates was guided by the *fimH*-types identification strategy we had developed, based on the analysis of single-nucleotide polymorphism of the FimH-f- and FimH-f-amplified sequenced sequences (Table 3).

Table 3. The strategy of identification of *fimH*-types during fimH locus genotyping of *E. coli* (summary table)

***fimH*-type**	***fimH* SNP(s)**	**GenBank A/N**	**N**
f-1	None (*E. coli* K-12 reference sequence)	GQ487190	151
f-2	C411G, G414A, T534C, C546T, C577T, A603G, T714C, A717G, G807A.	GQ486937	35
f-3	C411G, G414A, T534C, C546T, A603G, C654A, T714C, A717G, G807A.	GQ486930	41
f-4	C419T, T534C, A603G, G807A.	GQ486914	63
f-5	G420A, C546T, A603G, G807A.	GQ486919	42
f-6	T534C, C546T, T714A, A717G, G795A, G807A.	GQ486928	67
f-7	T714A, A717G, G807A.	GQ486915	69
f-8	C411G, G414A, T534C, C546T, T551C, A603G, C654A, T714C, A717G, G807A.	GQ487191	36
f-9	C411T, **G560A**, T714A, A717G, G807A.	GQ486920	62
f-10	C546T, **G645A**, G702C, T714A, A717G, G807A.	GQ487018	41
f-11	G807A.	GQ487032	25
f-12	C411G, G414A, T429A, T534C, A603G, T714C, A717G, G807A.	GQ486933	19
f-13	T534C, C546T, G550A, T714A, A717G, A751G, G795A, G807A.	GQ486923	8
f-14	C411G, G414A, T429A, T534C, C577T, A603G, C640T, G807A.	GQ486916	6
f-15	C489T, C639T, C668T, T714A, A717G, G807A.	GQ486948	42
f-16	T534C, A603G, G807A.	GQ487043	7
f-17	C411G, G414A, T450A, C546T, A603G, T714A, A717G, C788T, G807A.	GQ486952	12
f-18	C588T, T591C, A597G, A603G, G702C, T714A, A717G, **G741A**, G807A.	GQ487120	4
f-19	C432T, **C471T**, C489T, T534C, A603G, G702C, T714C, A717G, G807A.	GQ486932	5
f-20	**G412A**.	GQ486964	7
f-21	G420A, C546T, T714A, A717G, G807A.	GQ486982	3
f-22	T591A.	GQ487114	1
f-23	C489T, C546T, T714A, A717G, A751G, G795A, G807A.	GQ487129	3
f-24	C699T, T714A, A717G, G807A.	GQ487091	2
f-25	G414A, C419T, T450A, C546T, A603G, T714A, A717G, C788T, G807A.	GQ487063	2
f-26	C411G, G414A, T429A, T534C, C546T, C577T, A603G, T714C, A717G, G807A.	GQ487061	3
f-27	T450A, C546T, C588T, T591C, A597G, A603G, G702C, T714A, G807A.	GQ487078	10
f-28	C588T, T591C, A597G, A603G, A647T, T714A, A717G, **G794A**, G807A.	GQ487085	6
f-29	C411G, G414A, T429A, T534C, A603G, C640T, T714C, A717G, G807A.	GQ486924	8
f-30	C411G, G414A, T534C, C546T, T551C, A603G, C654A, **T707C**, T714C, A717G, G807A.	GQ486970	1
f-31	C411G, G414A, T534C, A603G, **G673C**, A717G, G807A.	GQ486999	1
f-32	**T446A**, C489T, C639T, C668T, T714A, A717G, G807A.	GQ487007	1
f-33	C432T, C489T, T534C, C588T, T591C, A597G, A603G, A647T, T714A, A717G, G807A.	GQ486959	1
f-34	T534C, C546T, G550A, **T618A**, T714A, A717G, A751G, G795A, G807A.	GQ487014	1
f-35	**A478T**.	GQ486938	3
f-36	C588T, T591C, A597G, A603G, A647T, G702C, T714A, A717G, G807A.	GQ487083	1
f-37	**T558C**, C699A, T714A, A717G, G795A, G807A.	GQ487051	2
f-38	C489T, C639T, C668T, **C693A**, T714A, A717G, G807A.	GQ487093	1

f-39	G408T, T450A, C546T, C588T, T591C, A597G, A603G, G702C, T714A, A792G, G807A.	GQ487109	2
f-40	**G479A**, A603G, **C657T**, T714A, A717G, G807A.	GQ487156	1
f-41	T534C, C546T, **G643A**, T714A, A717G, G795A, G807A.	GQ487138	1
f-42	G408T, C432T, C489T, T534C, C588T, T591C, A597G, A603G, A647T, T714A, A717G, G807A.	GQ487157	1
f-43	A603G, G702C, T714A, A792G, G807A.	GQ487155	6
f-44	G408T, C432T, C489T, T534C, T714A, A717G, **G738A**, G807A.	GQ487117	1
f-45	C411G, G414A, T534C, C577T, A603G, T714C, A717G, G807A.	GQ486894	5
f-46	C546T, A603G, T714A, A717G, G807A.	GQ486889	2
f-47	C411G, G414A, G442A, T534C, C546T, A603G, C654A, T714C, A717G, G807A.	KJ190188	2
f-48	C411G, G414A, T534C, C546T, C577T, A603G, **G632A**, T714C, A717G, G807A.	KJ190231	1
f-49	A603G, T714A, A717G, G807A.	KJ190226	3
f-50	C411G, G414A, T534C, C546T, A603G, T714C, A717G, G807A.	JQ658995	4
f-51	C411G, G414A, T534C, A603G, T714C, A717G, G807A.	FJ865725	30
f-52	C411G, G414A, G442A, T534C, C546T, A603G, **T641C**, C654A, T714C, A717G, G807A.	CP016497	2
The total number of analyzed nucleotide sequences of the *fimH* gene locus of *E. coli* representatives, published in GenBank NCBI			853

Note: N is the total number of nucleotide sequences of the *fimH* gene locus of *E. coli* representatives, published in GenBank NCBI. The GenBank A / N column indicates the identifiers of the reference sequences of *E. coli fimH* types. Additional SNPs not listed in previous studies are in gray. The type-specific SNPs of the individual *E. coli fimH* type are in bold and underlined.

The significant novelty of the modified strategy for identifying fimH-types in *fimH* locus genotyping *E. coli* is that the analysis of the corresponding nucleotide sequences of *E. coli* strains and isolates published in GenBank NCBI (f-50, f-51 and f-52). Also identified are type-specific SNPs specific to a single *fimH*-type of *E. coli*, as well as additional SNPs not mentioned in previous studies [9, 10].

Along with the summary table (Table 3), the strategy for identifying *fimH*-types in the *fimH* locus genotyping of *E. coli* is also framed and presented in the form of a generalized table for visual reference (Table 4).

The Tru26n Shi1o Shi1n Ef15o samples had a positive amplification signal at the *E. coli fimH* gene locus, which was confirmed by sequencing of a specific 878 bp-long PCR product (Fig. 2).

Table 4. The strategy of identification of *fimH*-types during fimH locus genotyping of *E. coli* (summary table)

f	Polymorphic nucleotide positions (SNP)																																																			
	408	411	412	414	419	420	429	432	442	446	450	471	478	479	489	534	546	550	551	558	560	577	578	591	597	603	618	632	639	640	641	643	645	647	654	657	668	673	693	699	702	707	714	717	738	741	751	788	792	794	795	807
1	G	C	G	G	C	G	T	C	G	T	T	C	A	G	C	T	C	G	T	T	G	C	C	T	A	A	T	G	C	C	T	G	G	A	C	C	C	G	C	C	G	T	T	A	G	G	A	C	A	G	G	G
2		G		A												C	T					T				G																	C	G								A
3		G		A												C	T									G									A								C	G								A
4					T											C										G																										A
5						A											T									G																										A
6																C	T																										A	G							A	A
7																																											A	G								A
8		G		A												C	T		C							G									A								C	G								A
9		T																			**A**																						A	G								A
10																	T																**A**								C		A	G								A
11																																																				A
12		G		A			A									C										G																	C	G								A
13																C	T	A																									A	G			G				A	A

ISSN: 1674-0440

14		G		A			A									C						T				G				T																						A
15															T														T								T						A	G								A
16																C										G																										A
17		G		A							A						T									G																	A	G				T				A
18																							T	C	G	G															C		A	G		**A**						A
19								T				**T**			T	C										G															C		C	G								A
20			**A**																																																	
21						A											T																										A	G								A
22																								A																												
23															T		T																										A	G			G				A	A
24																																								T			A	G								A
25				A	T						A						T									G																	A	G				T				A
26		G		A			A									C	T					T				G																	C	G								A
27											A						T						T	C	G	G															C		A									A
28																							T	C	G	G								T									A	G						**A**		A
29		G		A			A									C										G				T													C	G								A
30		G		A												C	T		C							G									A							**C**	C	G								A
31		G		A												C										G												**C**						G								A
32										**A**					T														T								T						A	G								A
33								T							T	C							T	C	G	G								T									A	G								A
34																C	T	A									**A**																A	G			G				A	A
35													**T**																																							
36																							T	C	G	G								T							C		A	G								A
37																				**C**																				A			A	G							A	A
38															T														T								T		**A**				A	G								A
39	T										A						T						T	C	G	G															C		A						G			A
40														**A**												G										**T**							A	G								A
41																C	T															**A**											A	G							A	A
42	T							T							T	C							T	C	G	G								T									A	G								A
43																										G															C		A						G			A

44	T							T							T	C																											A	G	**A**							A
45		G		A												C						T				G																	C	G								A
46																	T									G																	A	G								A
47		G		A					A							C	T									G									A								C	G								A
48		G		A												C	T					T				G		**A**															C	G								A
49																										G																	A	G								A
50		G		A												C	T									G																	C	G								A
51		G		A												C										G																	C	G								A
52		G		A					A							C	T									G					**C**				A								C	G								A

Note: f - *fimH*-type of *E. coli*. f-1 is the first *fimH* type represented by the reference sequence of the *E. coli* strain K-12 (GenBank A/N GQ487190). Additional SNPs not listed in previous studies are in gray. The type-specific SNPs of the individual E. coli fimH type are in bold and underlined.

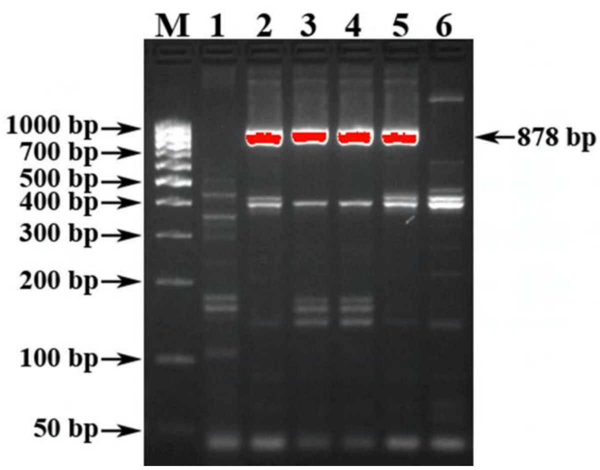

Fig. 2. Electrophoregram of the resulting PCR with FimH-f and FimH-r primers for *fimH* locus genotyping of *E. coli*
Designations: 1) Tru26o; 2) Tru26n; 3) Shi1o; 4) Shi1n; 5) Ef15o; 6) Ef15n.

The analysis of the sequences of the *E. coli fimH* gene that we sequenced by alignment with the corresponding nucleotide sequences previously deposited in the GenBank NCBI established the affiliation of the Tru26n sample of *E. coli* isolate with the type 7 (f-7), Shi1o и Shi1n – type 3 (f-3), and Ef15o – type 4 (f-4), respectively.
In addition, each of the sequences of the *E. coli fimH* gene locus that we sequenced was 100% homologous with the corresponding reference sequence, whose identifiers are also presented in Table 3.
It should be noted that earlier molecular analysis of the studied microbial cultures revealed a number of isolates to be producers of carcinogenic toxins, such as colibactin (*E. coli* monoculture isolates of Shi1o and Shi1n samples, isolates of mixed cultures of *E.coli* and *K. pneumonia* of Tru26o and Ef15n) and cytotoxic necrotizing factor 1 (Shi1o and Shi1n), which are pathogenetic tumor markers [13].
Consideration of the *fimH* types of *E. coli* in terms of potential tumor markers of colorectal cancer seems to be a relevant topic for further research [18, 19], since both the early and late phases of epithelial cell response to adhesion of *E. coli* demonstrate noticeable specificity for various *FimH+* strains [20] and can provide new data on molecular cascades initiated by bacteria in the process of tumor transformation.

4. Summary

1. Phylotyping of *E. coli* established that Shi1o and Shi1n samples belong to phylogroup B2, and Tru26o, Tru26n, Ef15o, and Ef15n samples belong to phylogroup A, which confirms the epidemiological association of identified isolates with the occurrence of colorectal cancer in the total research volume.
2. The *fimH* locus genotyping of *E. coli* established the affiliation of the Tru26n sample of *E. coli* isolate with the type 7 (f-7), Shi1o и Shi1n – type 3 (f-3), and Ef15o – type 4 (f-4), respectively.
3. An improved strategy for identifying *fimH*-types in *fimH* locus genotyping of *E. coli* provides for the presence of 52 *fimH*-types, three of which (f-50, f-51 and f-52) we substantiated in this paper.

4. Conclusions

In the cumulative scope of research, the phylotypic affiliation of *E. coli* isolates established by the set of genes confirms their epidemiological connection with the occurrence of colorectal cancer, and the established genotypic affiliation with an improved identification strategy for the *fimH* gene locus increases the level of knowledge about the genetic diversity of the analyzed microorganism.

Acknowledgements

This work was performed within the Russian Government Program of Competitive Growth of Kazan Federal University using equipment of Interdisciplinary Center of KFU (ID RFMEFI59414X0003) and supported by RFBR grant No 17-00-00060.

REFERENCES

[1] *Gao R. Gao Z, Huang L., Qin H.* Gut microbiota and colorectal cancer. Eur. J. Clin. Microbiol. Infect. Dis. 2017; 36 (5): 757-769. DOI: 10.1007/s10096-016-2881-8

[2] *Buc E., Dubois D., Sauvanet P., Raisch J., Delmas J., Darfeuille-Michaud A.* et al. High prevalence of mucosa-associated E. coli producing cyclomodulin and genotoxin in colon cancer. PLoS One. 2013; 8 (2): e56964. DOI: 10.1371/journal.pone.0056964

[3] *Gagnière J., Raisch J., Veziant J., Barnich N., Bonnet R., Buc E.* et al. Gut microbiota imbalance and colorectal cancer. World Journal of Gastroenterology. 2016; 22 (2): 501-518. DOI: 10.3748/wjg.v22.i2.501

[4] *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; 77 (11): 4696-4703.

[5] *Hennequin C., Forestier C.* OxyR, a LysR-type regulator involved in Klebsiella pneumonia mucosal and abiotic colonization. Infect. Immun. 2009; 77 (12): 5449-5457.

[6] *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; 9 (5). DOI: 10.1371/journal.pone.0096292

[7] *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; 20 (21): 6560-6572. DOI: 10.3748/wjg.v20.i21.6560

[8] *Clermont O., Christenson J.K., Denamur E., Gordon D.M.* The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environ. Microbiol. Rep. 2013; 5 (1): 58–65. DOI: 10.1111/1758-2229.12019

[9] *Dias R.C., Moreira B.M., Riley L.W.* Use of fimH single-nucleotide polymorphisms for strain typing of clinical isolates of Escherichia coli for epidemiologic investigation. J. Clin. Microbiol. 2010; 48 (2): 483-488. DOI: 10.1128/JCM.01858-09

[10] *Yun K.W., Kim D.S., Kim W., Lim I.S.* Molecular typing of uropathogenic Escherichia coli isolated from Korean children with urinary tract infection. Korean J. Pediatr. 2015; 58 (1): 20-27. DOI: 10.3345/kjp.2015.58.1.20

[11] *Abdallah K.S., Cao Y., Wei D.J.* Epidemiologic Investigation of Extra-intestinal pathogenic E. coli (ExPEC) based on PCR phylogenetic group and fimH single nucleotide polymorphisms (SNPs) in China. Int. J. Mol. Epidemiol. Genet. 2011; 2 (4): 339-353.

[12] *Magray M.S., Kumar A., Rawat A.K., Srivastava S.* Identification of Escherichia coli through analysis of 16S rRNA and 16S-23S rRNA internal transcribed spacer region sequences. Bioinformation. 2011; 6 (10): 370-371.

[13] *Nguyen N.T., Vafin R.R., Rzhanova I.V., Kolpakov A.I., Gataullin I.G., Tyulkin S.V.* et al. Molecular genetic analysis of microorganisms with intraepithelial invasion isolated from patients with colorectal cancer. Mol. Genet. Microbiol. Virol. 2016; 31 (1): 15-20. DOI: 10.3103/S0891416816010079

[14] *Clermont O., Bonacorsi S., Bingen, E.* Characterization of an anonymous molecular marker strongly linked to Escherichia coli strains causing neonatal meningitis. J. Clin. Microbiol. 2004; 42 (4): 1770-1772. DOI: 10.1128/JCM.42.4.1770-1772.2004

[15] *Clermont O., Bonacorsi S., Bingen E.* Rapid and simple determination of Escherichia coli phylogenetic group. Appl. Environ. Microbiol. 2000; 66 (10): 4555-4558. DOI: 10.1128/AEM.66.10.4555-4558.2000

[16] *Lescat M., Clermont O., Woerther P.L., Glodt J., Dion S., Skurnik D.* et al. Commensal Escherichia coli strains in Guiana reveal a high genetic diversity with hostdependant population structure. Environ. Microbiol. Rep. 2013; 5 (1): 49-57. DOI: 10.1111/j.1758-2229.2012.00374.x

[17] *Clermont O., Lescat M., O'Brien C.L., Gordon D.M., Tenaillon O., Denamur E.* Evidence for a human-specific Escherichia coli clone.

Environ. Microbiol. 2008; 10 (4): 1000-1006. DOI: 10.1111/j.1462-2920.2007.01520.x

[18] *Jin Y., Tang S., Li W., Ng S.C., Chan M.W., Sung J.J.* et al. Hemolytic E. coli Promotes Colonic Tumorigenesis in Females. Cancer Res. 2016; 76 (10): 2891-2900. DOI: 10.1158/0008-5472.CAN-15-2083

[19] *Zhao L., Zhang X., Zuo T., Yu J.* The Composition of Colonic Commensal Bacteria According to Anatomical Localization in Colorectal Cancer. Engineering. 2017; 3 (1): 90-97. DOI: 10.1016/J.ENG.2017.01.012

[20] *Mysorekar I.U., Mulvey M.A., Hultgren S.J., Gordon J.I.* Molecular regulation of urothelial renewal and host defenses during infection with uropathogenic Escherichia coli. J Biol Chem. 2002; 277 (9):7412-7419. DOI: 10.1074/jbc.M110560200