BRIEF REPORT

FREQUENCY OF SERINE PROTEASE AUTOTRANSPORTERS OF *Enterobacteriaceae* (SPATE) ENCODING GENES IN DIFFUSELY ADHERENT *Escherichia coli* (DAEC) ISOLATES FROM CHILDREN WITH AND WITHOUT DIARRHEA

Karen Ocampo[1,2,a], Maribel Riveros[1,3,b], Maria Pinedo-Bardales[1,4,c], Joaquim Ruiz[5,6,d], Theresa J. Ochoa[1,7,e]

[1] Laboratorio de Infectología Pediátrica, Instituto de Medicina Tropical "Alexander von Humboldt", Universidad Peruana Cayetano Heredia, Lima, Perú.
[2] Universidad Nacional Federico Villarreal, Lima, Perú.
[3] Facultad de Medicina, Universidad Peruana Cayetano Heredia, Lima, Perú
[4] Facultad de Ciencias, Universidad Nacional de Piura, Piura, Perú.
[5] ISGlobal, Hospital Clínic - Universitat de Barcelona, Barcelona, España.
[6] Laboratorio de Microbiología Molecular y Genómica Bacteriana, Universidad Científica del Sur, Lima, Perú.
[7] School of Public Health, University of Texas Health Science Center at Houston, Houston, Estados Unidos.
[a] Biologist; [b] Master in Molecular Biology; [c] Biology student; [d] Microbiologist, Doctor in Biology; [e] Physician specialized in Pediatric Infectious Diseases.

This study is part of the thesis of Ocampo K., Frecuencia de las proteínas Autotransportadoras Serin-Proteasa de *Enterobacteriaceae* (SPATE) en cepas de *Escherichia coli* difusamente adherente (DAEC) aisladas de niños con y sin diarrea [Bachelor Thesis]. Lima: Facultad de Ciencias Naturales y Matemática, Universidad Nacional Federico Villarreal; 2015.

ABSTRACT

The aim of the study was to determine the frequency of six genes encoding serine protease autotransporter proteins *Enterobacteriaceae* (SPATE) in diffusely adherent *Escherichia coli* (DAEC) isolates from children with (WD, n=63) and without diarrhea (WOD, n=41) from Lima, Peru. WOD were considered a control group. For the detection of the genes, 2 multiple PCRs were standardized: triple A (*sigA*, *pet*, *espP*) and triple B (*sat*, *pic*, *espC*). In both groups, the most frequent SPATE gene was *Sat* (39.7% of WD and 41.5% of WOD), followed by *spP* (20.6% and 9.7% in WD and WOD respectively). The other genes were detected in proportions lower than 10.0%, in the following order of frequency: *pet*, *sigA*, *espC* and *pic*, without significant differences between the groups. It was concluded that Sat is the most frequent SPATE in DAEC and that these strains may possess SPATE genes regardless of whether they are isolated in WD or WOD.

Keywords: Diffusely Adherent *Escherichia coli*; Virulence Factors; Diarrhea; Child (source: MeSH NLM).

Cite as: Ocampo K, Riveros M, Pinedo-Bardales M, Ruiz J, Ochoa TJ. Frecuencia de genes que codifican proteínas autotransportadoras serin-proteasa de *Enterobacteriaceae* (SPATE) en cepas de *Escherichia coli* difusamente adherente (DAEC) provenientes de niños con y sin diarrea. Rev Peru Med Exp Salud Publica. 2021;38(1):124-9. doi: https://doi.org/10.17843/rpmesp.2021.381.5750

Correspondence: Maribel Riveros; Universidad Peruana Cayetano Heredia, Honorio Delgado 430, San Martín de Porres, Lima, Perú; maribel.riveros@upch.pe

Received: 06/05/2020
Approved: 30/09/2020
Online: 24/02/2021

INTRODUCTION

Diarrhea is one of the main causes of death in children under five years of age in the world. In this age group, *Escherichia coli* is one of the most frequent etiological agents [(1)]. According to their pathogenesis, diarrheagenic *E. coli* (DEC) are divided into six pathotypes: enteroaggregative (EAEC), enterotoxigenic (ETEC), enteropathogenic (EPEC), verotoxigenic (VTEC), enteroinvasive (EIEC) and diffusely adherent (DAEC) [(2)].

DAEC is a heterogeneous group and has been particularly associated with persistent diarrhea in developing countries, mainly affecting children older than one year [(2)]. In Peru, DAEC have been found in samples of children under five years of age with persistent

https://doi.org/10.17843/rpmesp.2021.381.5750

diarrhea with frequencies of 15.0% and 17.0% for Lima and Lambayeque, respectively [3,4]. In addition, it has been reported that DAEC shows difference in adherence patterns, actin polymerization and motility depending on whether it is isolated from children with diarrhea or asymptomatic children, and the determination of virulence factors present in these strains is suggested [5].

Pathogenic strains of *E. coli*, both extraintestinal and DEC, are characterized by virulence factors, which help them overcome host defenses and colonize or invade the urinary tract (uropathogenic strains) or gastrointestinal tract (DEC strains) [6,7]. Virulence factors include toxins, adhesins and other hydrolytic enzymes secreted through pathways known as protein secretion systems [8]. These pathways are classified into five main groups: type I, II, III, IV and autotransporters. Within the group of autotransporter proteins, the most frequent subclass consists of the serine-protease autotransporters of *Enterobacteriaceae* (SPATE), which have been identified in gram-negative bacteria such as *Salmonella* and *Shigella* [9,10].

The most reported SPATE in DECs is the secreted autotransporter toxin (*sat*) [2,11,12]. Although the presence of SPATEs in DAEC strains has been scarcely studied, the *sat* has been reported to be the most frequent, being found in a higher percentage of strains isolated from patients with diarrhea than in those from asymptomatic carriers [12]. The scarce scientific evidence on the main SPATEs present in DAEC strains makes it relevant to analyze their frequency in this pathotype. Therefore, the aim of the study was to determine the frequency of six SPATE-encoding genes in DAEC isolates from children with and without diarrhea.

KEY MESSAGES

Motivation for conducting the study: It is necessary to evaluate the presence of Serine protease autotransporters of *Enterobacteriaceae* (SPATE) proteins in diffusely adherent *Escherichia coli* (DAEC) strains, because these proteins facilitate host invasion during bacterial infection.

Main findings: Genes encoding SPATE can be found in DAEC strains, both in samples from patients with diarrhea and in healthy children. The sat gene is the most frequent SPATE in these strains.

Implications: Knowing the genes that encode the molecular mechanisms involved in the interaction between humans and the infectious agent will help guiding research aimed at diagnostic and control tests or the development of future vaccine candidates.

THE STUDY

Strains

We analyzed 104 strains isolated from children under five years old with diarrhea (n=63) and asymptomatic children, which were considered as a control group (n=41). The isolates came from three cohort studies conducted in the district of Lima (Peru) between 2006 and 2010, being molecularly identified as DAEC by detection of the *daaD* gene [13] and reported by Ochoa *et al.* [3]. The DAEC strains were reactivated on MacConkey agar and incubated for 24 h at 37 °C. Subsequently, DNA was extracted by the heat shock method.

The strains included in this study were obtained from studies (SIDISI codes: 50744, 57025 and 50908) that were approved by the Ethics Committee of the Universidad Peruana Cayetano Heredia.

SPATE detection

The presence of SPATE was determined by two multiplex polymerase chain reactions (PCR) using the primers reported in Table 1. Initially, single PCRs were standardized for the individual detection of each gene (Table 1) to establish the optimal conditions.

After evaluating the effectiveness of the amplification of the products at appropriate temperatures, we proceeded to the design of multiple PCRs and the selection of the best primer combination. For this, we used the Integrated DNA Technologies website (http://www.idtdna.com), where the guanine and cytokine content (% GC), melting temperature (Tm), hairpin formation and generated dimers were evaluated. Subsequently, the formation of secondary structures with the target DNA was analyzed in *mfold* (http://mfold.rna.albany.edu/?q=mfold/dna-folding-from).

Specific, defined and distinguishable bands (Figure 1A and 1B) were obtained by multiple amplification of *pet*, *sigA* and *espP* (triplex A), and *sat*, *espC* and *pic* (triplex B).

To standardize triplex A, the hybridization temperature gradient was evaluated from 54 to 64 °C at a final volume of 25 µL containing: dNTP (0.23 mM), MgCl2 (1.7 µM), *pet* and *sigA* primers at 0.5 µM, *espP* primers at 0.6 µM, Taq buffer (final concentration 1X) and GoTaq DNA Polymerase (Promega, France) (5U/µM); under the following conditions:

Table 1. List of primers used in our study.

Technique	Gene	Primer	Base pairs	Reference
Triplex A PCR	*espP*	Forward: GTCCATGCAGGGACATGCCA Reverse: TCACATCAGCACCGTTCTCTAT	547	9
	pet	Forward: GGCACAGAAT AAAGGGGTGTTT Reverse: CCTCTTGTTTCCACGACATAC	302	9
	sigA	Forward: CCGACTTCTCACTTTCTCCCG Reverse: CCATCCAGCTGCATAGTGTTTG	430	9
Triplex B PCR	*sat*	Forward: TCAGAAGCTCAGCGAATCATTG Reverse: CCATTATCACCAGTAAAACGCACC	930	9
	espC	Forward: TAGTGCAGTGCAGAAAGCAGTT Reverse: AGTTTTCCTGTTGCTGTATGCC	301	8
	pic	Forward: ACTGGATCTTAAGGCTCAGGAT Reverse: GACTTAATGTCACTGTTCAGCG	572	9

PCR: Polymerase Chain Reaction

initial denaturation at 94 °C for five minutes and 30 cycles with denaturation at 94 °C for one minute, hybridization for 30 seconds, elongation at 72 °C for one minute, and a final elongation at 72 °C for five minutes. Under these conditions, triplex A showed better band definition at the 60 °C hybridization temperature.

In the standardization of triplex B, we evaluated hybridization temperatures from 54 to 64 °C, at a final volume of 25 µL containing: dNTP (0.08 mM), MgCl2 (1.7 µM), primers (*espC*, *pic* and *sat*) at 0.5 µM, Taq buffer (final concentration 1X) and Taq (5U/µM); under the following conditions: initial denaturation at 94 °C for five minutes and 30 cycles with denaturation at 95°C for one minute, hybridization for one minute, elongation at 72 °C for one minute, and a final elongation at 72 °C for ten minutes. Higher definition of PCR products was observed at the 64 °C temperature.

Reactions were conducted on the Applied Biosystems Veriti 96-Well Thermal Cycler (Fisher Scientific, USA). The negative control was sterile water and the positive controls were *E. coli* strains EAEC 042 for *pet* and *pic*, EHEC 933 for *espP* and EPEC 2348/64 for *espC*, and *Shigella flexneri* strain 1106 for *sigA* and *sat*. The control strains were obtained from the internal

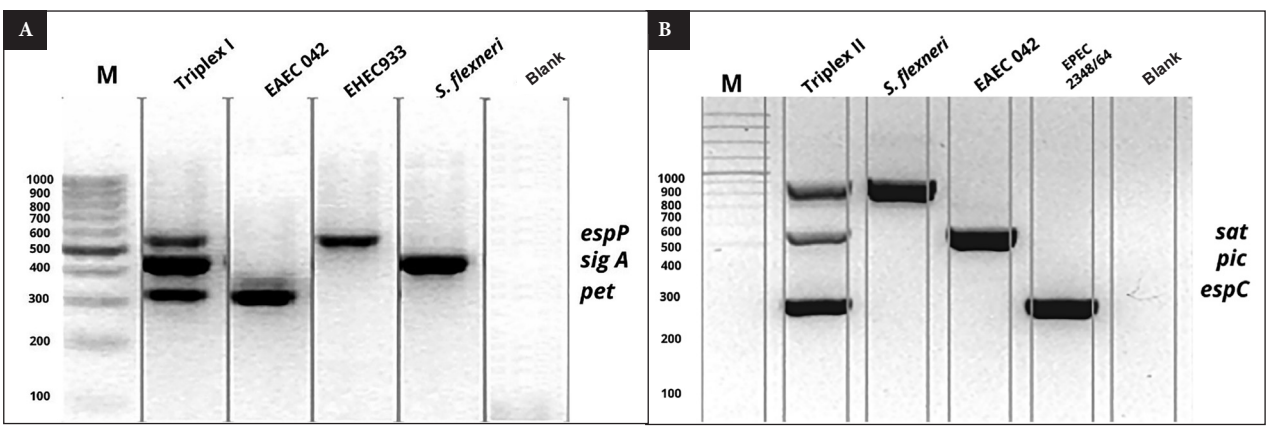

Figure 1. A. 2% agarose gel, second lane corresponds to triplex A, lanes 3, 4 and 5 correspond to individual genes, *espP* (547 bp), *sigA* (430 bp), *pet* (302 bp), hybridization temperature of 60 °C, blank (negative control). **B.** 2% agarose gel, second lane corresponds to triplex B, lanes 3, 4 and 5 corresponds to individual genes, *sat* (930 bp), *pic* (572 bp), *espC* (320 bp), hybridization temperature of 64 °C, blank (negative control).

https://doi.org/10.17843/rpmesp.2021.381.5750

Table 2. Frequencies of SPATE genes in DAEC strains from children with and without diarrhea under one year of age.

Characteristic	SPATE genes					
	pet	*sigA*	*espP*	*sat*	*pic*	*espC*
CWD (n=63)	5 (7.9%)	5 (7.9%)	13 (20.6%)	25 (39.7%)	0	0
CWOD (n=41)	2 (4.9%)	5 (12.2%)	4 (9.7%)	17 (41.5%)	2 (4.9%)	2 (4.9%)

CWD: children with diarrhea; CWOD: children without diarrhea.

collection of the Laboratory of Pediatric Infectious Diseases of the "Alexander von Humboldt" Institute of Tropical Medicine.

PCR products were visualized on 2% agarose gels, stained with ethidium bromide at 90V for 80 minutes.

Data analysis

The frequency and comparison of all variables were obtained by the Chi-square test using the Epi-Info 7 program. The GenAlEx program was used for the analysis of genomic profiles.

FINDINGS

Of the 104 isolates included in the study, 58.6% had at least one gene coding for SPATEs.

The most frequent gene was *sat* (40.4%), found in similar proportions in isolates from children with diarrhea (39.7%) and asymptomatic children (41.5%); followed by *espP* (16.3%), *sigA* (10.0%) and *pet* (7.0%). The *pic* and *espC* genes had a frequency of 1.9%. The *espP* gene was found to be associated with 20.6% of the samples from diarrheic cases; on the other hand, it was detected in 9.7% of the controls (isolates from asymptomatic children), although without significant differences (Table 2).

On the other hand, no association was observed between the presence of the remaining SPATEs and the origin (either case or control) of the DAEC strains.

Twelve genetic profiles were identified, determined by the combinations of genes coding for SPATEs (Table 3). The most common profile among the diarrheal samples was the absence of SPATEs (41.4%), followed by the presence of only the *sat* gene (30.8%). The remaining nine profiles were less frequent with none exceeding ten isolates. None of the profiles was associated with the presence or absence of diarrhea.

DISCUSSION

In the samples analyzed, the *sat* gene was found to have the highest frequency, both in cases (39.7%) and in controls (41.5%), a result similar to the report of Taddei *et al.* [11], who found 44.0% positivity for this gene after analyzing 146 DAEC strains. Similarly, Boisen *et al.* [9] detected the *sat* gene in 30.0% of DAEC strains analyzed. Although, unlike that reported by other authors [12], the frequency of *sat* was

Table 3. Genetic profile of DAEC strains isolated from children with and without diarrhea.

Profiles	*pet*	*espP*	*sigA*	*sat*	*pic*	*espC*	CWD (n=63)	CWOD (n=41)
1							26	17
2				+			20	12
3		+					6	2
4		+		+			4	1
5	+		+				2	0
6			+				2	2
7	+		+			+	1	0
8	+	+		+			1	0
9	+	+					1	1
10	+			+			0	1
11			+	+			0	3
12						+	0	2

CWD: children with diarrhea; CWOD: children without diarrhea; +: presence of the gene.

similar in cases and controls. Mansan *et al.* observed that the expression of the *sat* gene is related to lesions in epithelial cell tight junctions, leading to increased permeability and fluid loss, and that the presence of this gene was associated with cases of diarrhea (46.0% vs. 19.0%; p value <0.05) [14]. In our study, this association was not found; however, the results are in accordance with the research by Toloza *et al.* [15], where they concluded that the role of *sat* as a virulence factor depends on the genetic load of the strains. In this regard, the possibility of alterations at the gene expression level, point mutations that affect its functionality and bacterial load must be taken into account, which in the case of other DEC strains (EPEC) has been directly related to the presence/absence of symptoms [16].

The second most frequent gene was *espP*, which is strongly associated with atypical EHEC and EPEC, being reported less frequently in other DEC, as in the case of typical EAEC or EPEC. The *espP* gene encodes a protein capable of degrading coagulation factor V, thus participating in the production of uremic-hemolytic syndrome, which is characteristic of EHEC. It is also responsible for disrupting complement factor C3/C3b, which facilitates colonization of the intestine [10].

The *sigA* gene had a frequency of 10.0%, similar to that described by Boisen *et al.* [9]. This gene was initially reported in the SHI-1 pathogenicity island found in *S. flexneri* 2a [17]; however, the detection of this gene in the absence of other components of this pathogenicity island suggests it could be present in other genetic structures [9].

The remaining three genes were detected in less than 10% of the samples. The *pet* gene was detected in 7.0% of the strains, mostly from children with diarrhea. In other DECs, such as EPEC and EAEC, *pet* has been reported with frequencies between 5 and 20.0%, respectively [7,9,18]. The *pet* gene is capable of producing enterotoxigenic and cytopathic effects, and has the ability to degrade phodrin, which produces the disorganization of the actin of the cellular cytoskeleton [19].

Both *pic* and *espC* were found with frequencies lower than 2.0% and only in DAEC strains from asymptomatic children. *espC* has been previously described in EPEC strains [17]. The protein encoding this gene plays a relevant role in cell death induced by EPEC, regulating the secretion of virulence factors such as EspA or EspD. *pic* has been reported in other DECs, such as EAEC and UPEC [20].

Analysis of the concomitant presence of different SPATE showed 12 different gene association profiles. However, none of the observed patterns could be related to cases or controls. This non-association may be due to several factors, including the presence of different allelic variants or the expression levels of the genes evaluated [16].

A limitation of this study is that, due to the absence of clinical data from the patients, it was not possible to perform a more complex analysis or to relate it to other symptoms. Also, the number of isolates available may have influenced the failure to detect a significant difference between the case and control groups. Nevertheless, this study presents the first description of SPATEs genes in DAEC strains isolated from Peruvian children.

In conclusion, 58.6% of the DAEC strains isolated from children with and without diarrhea had at least one SPATE-encoding gene. In both groups, *sat* and *espP* were the most frequent genes. Overall, our results showed heterogeneity at the level of SPATEs in DAEC strains, so further studies with a larger number of isolates are needed, both from children with diarrhea and asymptomatic children. The detection of SPATEs is important because these proteins facilitate host invasion during bacterial infection and some coding genes (such as *pet* and *espP*) can be transmitted by plasmids to other bacterial species.

Authorship contributions: KO and MR participated in the collection of results, analysis and interpretation of data, and writing of the article. MP performed the analysis and interpretation of data and writing of the article. JR and TJO participated in the conception and design of the article, by obtaining funding, writing the article, and carrying out a critical revision. All authors approved the final version and assume responsibility for its content.

Funding: This research was supported by the Spanish Agency for International Cooperation and Development (AECID) (D/030509/10). Joaquim Ruiz was funded by the I3 program of the Instituto de Salud Carlos III (grant number: CES11/012) and by the Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica (FONDECYT - Peru) under the "Proyecto de Mejoramiento y Ampliación de los Servicios del Sistema Nacional de Ciencia, Tecnología e Innovación Tecnológica" (contract number 08-2019-FONDECYT-BM-INC-INV).

Acknowledgements: ISGlobal is a member of the CERCA program, Generalitat de Catalunya.

Conflicts of interest: The authors declare no conflicts of interest.

Supplementary material: Available in the electronic version of the RPMESP.

REFERENCES

1. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, *et al.* Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. Lancet. 2013;382:209–22. doi: 10.1016/S0140-6736(13)60844-2.
2. Servin AL. Pathogenesis of human diffusely adhering Escherichia coli expressing Afa/Dr adhesins (Afa/Dr DAEC): Current insights and future challenges. Clin Microbiol Rev. 2014;27:823–69. doi: 10.1128/ CMR.00036-14.
3. Ochoa TJ, Mercado EH, Durand D, Rivera FP, Mosquito S, Contreras C, *et al.* Frecuencia y patotipos de Escherichia coli diarrogénica en

 https://doi.org/10.17843/rpmesp.2021.381.5750

niños peruanos con y sin diarrea. Rev Peru Med Exp Salud Publica. 2011;28(1):13-20. doi: 10.1590/s1726-46342011000100003.
4. Yacarini-Martínez AE, Arriaga-Deza EV, Alvarado-Pineda RL, Fupuy-Chung JA. Genes de virulencia de Escherichia coli detectados en muestras diarreicas de niños de la Región Lambayeque - Perú. Horiz Med (Lima). 2019;19(1):7-12. doi: 10.24265/horizmed.2019.v19n1.02.
5. Riveros M, Barletta F, Cabello M, Durand D, Mercado EH, Contreras C, *et al*. Patrones de adherencia de cepas de Escherichia coli Difusamente adherente (DAEC) provenientes de niños con y sin diarrea. Rev Peru Med Exp Salud Publica. 2011;28(1):21-8. doi: 10.1590/s1726-46342011000100004.
6. Johnson JR. Virulence factors in Escherichia coli urinary tract infection. Clin Microbiol Rev. 1991;4:80–128. doi: 10.1128/cmr.4.1.80.
7. Durand D, Contreras CA, Mosquito S, Ruíz J, Cleary TG, Ochoa TJ. pic gene of enteroaggregative Escherichia coli and its association with diarrhea in Peruvian children. Pathog Dis. 2016;74:ftw054. doi: 10.1093/femspd/ftw054.
8. Restieri C, Garriss G, Locas M-C, Dozois CM. Autotransporter-encoding sequences are phylogenetically distributed among Escherichia coli clinical isolates and reference strains. Appl Environ Microbiol. 2007;73(5):1553– 62. doi: 10.1128/AEM.01542-06.
9. Boisen N, Ruiz-Perez F, Scheutz F, Krogfelt KA, Nataro JP. High prevalence of serine protease autotransporter cytotoxins among strains of enteroaggregative Escherichia coli. Am J Trop Med Hyg. 2009;80(2):294–301. doi: 10.4269/ajtmh.2009.80.294.
10. Weiss A, Brockmeyer J. Prevalence, biogenesis, and functionality of the serine protease autotransporter. EspP Toxins (Basel). 2012;5(1):25-48. doi: 10.3390/toxins5010025.
11. Taddei CR, Moreno ACR, Fernandes Filho A, Montemor LP., Martinez MB. Prevalence of secreted autotransporter toxin gene among diffusely adhering Escherichia coli isolated from stools of children. FEMS Microbiol Lett. 2003;227:249–53. doi: 10.1016/S0378-1097(03)00688-8.
12. Guignot J, Chaplais C, Coconnier-Polter MH, Servin AL. The secreted autotransporter toxin, Sat, functions as a virulence factor in Afa/Dr diffusely adhering Escherichia coli by promoting lesions in tight junction of polarized epithelial cells. Cell Microbiol. 2007;9(1):204-21. doi: 10.1111/j.1462-5822.2006.00782.x.
13. Guion CE, Ochoa TJ, Walker CM, Barletta F, Cleary TG. Detection of Diarrheagenic Escherichia coli by use of melting-curve analysis and real-time multiplex PCR. J Clin Microbiol. 2008;46:1752–7. doi: 10.1128/ JCM.02341-07.
14. Mansan-Almeida R, Pereira A, Giugliano L. Diffusely adherent Escherichia coli strains isolated from children and adults constitute two different populations. BMC Microbiol. 2013;13:22. doi: 10.1186/1471-2180-13-22.
15. Toloza L, Giménez R, Fábrega MJ, Alvarez CS, Aguilera L, Cañas MA, *et al*. The secreted autotransporter toxin (Sat) does not act as a virulence factor in the probiotic Escherichia coli strain Nissle 1917. BMC Microbiol. 2015;15:250. doi: 10.1186/s12866-015-0591-5.
16. Barletta F, Ochoa TJ, Cleary TG. Multiplex Real-Time PCR (MRT-PCR) for diarrheagenic. Methods Mol Biol. 2013;943: 307-14. doi: 10.1093/cid/cir730.
17. Al-Hasani K, Henderson IR, Sakellaris H, Rajakumar K, Grant T, Nataro JP, *et al*. The sigA gene which is borne on the she pathogenicity island of Shigella flexneri 2a encodes an exported cytopathic protease involved in intestinal fluid accumulation. Infect Immun. 2000;68(5):2457–63. doi: 10.1128/iai.68.5.2457-2463.2000.
18. Abreu AG, Bueris V, Porangaba TM, Sircili MP, Navarro-Garcia F, Elias WP. Autotransporter protein-encoding genes of diarrheagenic Escherichia coli are found in both typical and atypical enteropathogenic E. coli strains. Appl Environ Microbiol. 2013;79(1):411–4. doi: 10.1128/AEM.02635-12
19. Navarro-Garcia F, Sonnested M, Teter K. Host-Toxin interactions involving EspC and Pet, two serine protease autotransporters of the Enterobacteriaceae. Toxins (Basel). 2010;2(5):1134–47. doi: 10.3390/toxins2051134.
20. Pokharel P, Habouria H, Bessaiah H, Dozois CM. Serine protease autotransporters of the Enterobacteriaceae (SPATEs): out and about and chopping it up. Microorganisms. 2019;7(12):594. doi: 10.3390/microorganisms7120594.