RESEARCH LETTER – Food Microbiology

Variants of astA gene among extra-intestinal Escherichia coli of human and avian origin

Renato Pariz Maluta¹, Janaína Luisa Leite¹, Thaís Cabrera Galvão Rojas¹, Isabel Cristina Affonso Scaletsky², Elisabete Aparecida Lopes Guastalli³, Marcelo de Carvalho Ramos⁴ and Wanderley Dias da Silveira¹,*

¹Bacterial Molecular Biology Laboratory, Department of Genetics, Evolution and Bioagents, Institute of Biology, State University of Campinas (UNICAMP), CEP 13083-875, Campinas, SP, Brazil, ²Department of Microbiology, Department of Immunobiology and Parasitology, Federal University of São Paulo (UNIFESP), CEP 04021-001, São Paulo, SP, Brazil, ³Instituto Biológico, CAPTAA, Unidade de Pesquisa e Desenvolvimento de Bastos, CEP 17690-000, Bastos, SP, Brazil and ⁴Department of Internal Medicine, School of Medicine, State University of Campinas (UNICAMP), CEP 13083-894, Campinas, SP, Brazil

*Corresponding author: Bacterial Molecular Biology Laboratory, Department of Genetics, Evolution and Bioagents, Institute of Biology, State University of Campinas (UNICAMP), Rua Candido Rondon, 390, Barração da Genética, CEP 13083-875, Campinas, SP, Brazil. Tel: +55-19-35211160; E-mail: wds@unicamp.br

One sentence summary: Sequencing and expression of an enterotoxin in Escherichia coli associated with extraintestinal diseases.

Editor: Richard Calendar

ABSTRACT

Many Escherichia coli strains harbour astA, which is the gene encoding the enteroaggregative E. coli heat-stable enterotoxin (EAST1). This gene is embedded in a putative transposase (ORF1) and presents polymorphism in diarrheagenic strains. Although astA and orf1 are detected in extraintestinal strains, little is known about polymorphism and differential gene transcription in this pathotype. In the present work, extraintestinal E. coli from humans (ExPEC - Extraintestinal Pathogenic E. coli) and poultry (APEC - Avian Pathogenic E. coli) were assayed to verify the presence of astA/orf1 and possible polymorphisms in these genes. Three astA/orf1 patterns were detected via Sanger sequencing. Pattern 1 was novel and represented an astA pseudogene. Pattern 2 and pattern 3 presented distinct amino acids within the reading frame encoding astA and were identical to the sequences found in EAEC 17-2 and EAEC 042, respectively. Regarding the frame encoding ORF1, all mutations detected in the three patterns were neutral. The transcripts of astA/orf1 in vitro were underregulated in strains possessing the pattern 1 sequence. The results demonstrate that the same astA sequences may be detected in diarrheagenic and extra-intestinal E. coli. However, extraintestinal isolates may also present an astA pseudogene that has not been reported in diarrheagenic E. coli.

Keywords: Escherichia coli; APEC; ExPEC; EAST1; transposase

INTRODUCTION

The enteroaggregative Escherichia coli heat-stable enterotoxin (EAST1) was first observed in enteroaggregative E. coli (EAEC), using the Ussing chamber model (Savarino et al. 1991). Later, this toxin (or its encoding gene astA) was found in other E. coli pathotypes, including enterotoxigenic E. coli (ETEC), enteropathogenic...
E. coli (EPEC) (Ménard and Dubreuil 2002) and other strains isolated from humans or animals (Yamamoto and Nakazawa 1997; Yamamoto and Taneike 2000; Maluta et al. 2014a,b,c,d). The toxin EAST-1 was detected in E. coli associated with outbreaks. In the USA, an O39 E. coli lacking heat-labile (LT) and heat-stable (STa, STb), but possessing the locus of enterocyte effacement (LEE) and EAST-1 was isolated from ill patrons in a restaurant (Hedberg et al. 1997). In Japan, a waterborne diarrhoea outbreak was associated with E. coli possessing the gene eae (an LEE component) and astA (Yatsuyanagi et al. 2003). The previous works reported the detection of astA in E. coli associated with outbreaks show that this gene might confer the ability to produce diarrhoea by E. coli.

Besides the original sequence of astA detected in EAEC strain 17-2, different sequences were found in strain EAECh4 (Yamamoto et al. 1997) and a subgroup of atypical EPEC (Silva et al. 2014). Furthermore, the analysis of a 1.3-kb DNA sequence of a recombinant plasmid (pAM1) showed that, besides astA, the same sequence encoded a putative transposase (ORF1) in the same coding region, but in a different reading frame (McVeigh et al. 2000).

Although a number of astA sequences of diarrheagenic E. coli have been reported, little is known about the sequences in extraintestinal strains. Thus, this study was designed to characterise astA in extraintestinal E. coli.

**MATERIALS AND METHODS**

*Escherichia coli* strains

Sixteen *E. coli* strains, known to harbour astA, both by PCR and colony hybridisation (Maluta et al. 2014c), were tested in this work. They belong to the bacterial collection of the Bacterial Molecular Biology Laboratory of Institute of Biology from State University of Campinas (UNICAMP), Campinas, SP, Brazil. Twelve strains were isolated of extraintestinal diseases from poultry (APEC) and four strains from humans (ExPEC). They were characterised regarding their serotype, pulse-field gel electrophoresis (PFGE) profile and sequence type (ST) in a previous work (Maluta et al. 2014c). Only strains with distinct PFGE patterns were used in the present work. Six APEC isolates were O-typable and belonged to five distinct serogroups (O2, O5, O6, O8 and O9). Two ExPEC isolates were O-typable and presented the serogroups O15 or O19. All isolates presented distinct STs, except two APEC isolates belonging to ST 117. Further details are shown in Table 1.

**Detection of astA variants by DNA sequencing**

Individual *E. coli* isolates were cultured in LB medium at 37°C overnight. These cultures were used for the preparation of DNA templates using a thermal lysis procedure (http://www.apzec.ca/en/APZEC/Protocols/pdfs/ECL-PCR_Protocol.pdf). PCR was performed as described previously (Maluta et al. 2014b) using primers EAST13a (5'-GCCCAGTCGACATCCCAGTATGC) and EAST12b (5'-CTGCAGTCGCGTTCTCCTCTCGT) (Yamamoto and Echeverria 1996). DNA sequencing was performed at 'Centro de Biologia Molecular e Engenharia Genética (CBMEG - UNICAMP)'. For sequencing, amplicons were purified with a column-based kit (Pure Link Quick PCR Purification Kit, Invitrogen, USA). The purified product was sequenced using Sanger methodology using an ABI PRISM 3700 DNA Analyzer (Applied Biosystems with the program PCR-BD3700).

**Table 1. Serotypes, ST types and levels of astA/orf1 transcripts among APEC and human ExPEC. Level of transcripts are presented as fold changes compared to the prototype EAEC 042 levels. Error bars represent the standard deviations (SD) of the ∆ΔCt value.**

<table>
<thead>
<tr>
<th>Pattern Strain ID Category</th>
<th>Serotype ST (Cplx)</th>
<th>Mean ± SD</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EC200 APEC O8:H9</td>
<td>88 (23)</td>
<td>0.486 ± 0.087</td>
</tr>
<tr>
<td>2</td>
<td>EC113 APEC O2:H5</td>
<td>746</td>
<td>0.401 ± 0.085</td>
</tr>
<tr>
<td>3</td>
<td>EC125 APEC O4:HNT</td>
<td>57 (350)</td>
<td>1.040 ± 0.005</td>
</tr>
<tr>
<td>4</td>
<td>UEL181 APEC O8:H9</td>
<td>117</td>
<td>0.583 ± 0.227</td>
</tr>
<tr>
<td>5</td>
<td>UEL175 APEC O4:HNT</td>
<td>117</td>
<td>1.115 ± 0.309</td>
</tr>
<tr>
<td>6</td>
<td>334/EC APEC</td>
<td>O9:H4</td>
<td>1711</td>
</tr>
<tr>
<td>7</td>
<td>H276/12 ExPEC</td>
<td>ONTH10</td>
<td>10 (10)</td>
</tr>
<tr>
<td>8</td>
<td>32/EC APEC</td>
<td>O4:HNT</td>
<td>4133</td>
</tr>
</tbody>
</table>

1Isolated from blood culture.

2Isolated from urinary tract infection.

**Quantitative real-time RT-PCR**

Bacterial cultures were grown overnight aerobically in LB medium at 37°C until reaching an OD400 = 0.6. RNA from three biological samples was extracted by using the RNAeasy Mini Kit (Qiagen, Germany). Primers RT_astA_1 (5’-GGCCGGATCCAGTATGC) and RT_astA_2 (5’-GTCCTTCCATGACGAGAACG) used for the real-time assays were constructed by using Primer Express v1.5 software (Applied Biosystems). The quantitative real-time PCR (qRT-PCR) was performed as described previously (de Pace et al. 2010). Gene rpoA was used as the endogenous control. Data collection was performed by using ABI Sequence Detection 1.3 software (Applied Biosystems). Data were normalised to levels of rpoA and analysed by using the comparative critical threshold (CT) method. Statistical significance was determined by the Student’s t-test.

**Protein modelling**

Protein modelling was done with RaptorX (Källberg et al. 2012).

**Nucleotide sequence accession numbers**

The nucleotide sequences of the astA patterns have been entered into the GenBank databases under the accession numbers KJ149553-68.

**RESULTS**

The sequencing of the amplicon, generated after PCR with primers EAST13a and EAST12b, yielded three different patterns, when compared to the astA sequence of the prototype strain EAEC 042 (Yamamoto and Taneike 2000) (Fig. 1). When compared with the sequence of the putative transposase (ORF1) of pAM1 (McVeigh et al. 2000), three patterns were also noticed (-1 reading frame with reference with astA).

**References**

- Yamamoto and Nakazawa (1997)
- Yamamoto and Taneike (2000)
- Maluta et al. (2014a,b,c,d)
Figure 1. Alignment of sequences obtained by using the primers EAST13a and EAST13b. Pattern 1, which was detected in APEC, is novel and represents an astA (EAST1) pseudogene (GenBank accession number KJ149556). Pattern 2, found in APEC and human ExPEC, is identical to the sequences in pAPEC-O2-colV (GenBank accession number AV545598) and EAEC 17-2 (Savarino et al. 1991). Pattern 3, detected in APEC and human ExPEC, is identical to the sequences found in the prototype strain EAEC 042 (Yamamoto et al. 1997) and pAM1 (GenBank accession number AF143819). The three patterns presented the same predicted ORF1 (partial) when compared with the sequence of pAM1. Dots represent identical nucleotides or predicted amino acids residues when compared with pAM1. Stars indicate a stop codon.

Two APEC strains presented a nonsense mutation at the 16th EAST1 codon (TG→TA) and a missense one at the 237th nucleotide (C→T). The nonsense mutation altered the product (Ser→STOP), predicting a truncated EAST1 protein (Fig. 1). The partial ORF1 sequence contained a synonymous substitution at the 283rd codon (GCT→GCC). Interestingly, these APEC strains presented different serotypes and STs (Table 1).

Six APEC and one human ExPEC shared the same astA pattern, which is identical to the sequences previously detected in the plasmid colV from the APEC O2 (O2:K2) (Johnson et al. 2006) and EAEC 17-2 (Savarino et al. 1991). This sequence presented a missense mutation at the 33rd EAST1 codon (AGG→AGG), generating a predicted peptide with a different amino acid (Lys→Arg) (Fig. 1). The predicted model of the protein indicates the presence of a secondary structure with 57% of helix and 42% of loops (Fig. 2). At the 282nd codon, the partial ORF1 sequence contained a synonymous substitution (CAA→CAA). Two APEC strains with this pattern shared the serotype ONT:H4 and ST117 (Table 1).

Four APEC and three ExPEC strains presented the same astA (EAST1) pattern found in the prototype strain EAEC 042. Similar to pattern 2, the predicted model of the protein indicates the presence of a secondary structure with 57% of helix and 42% of loops (Fig. 2). They also presented the same EAST1 and ORF1 pattern found in pAM1 (Fig. 1). These strains were diverse, except one APEC and one ExPEC strain that belonged to ST1011 (Table 1).

Thirteen strains were chosen for the qRT-PCR assays (pattern 1, n = 2; pattern 2, n = 5 and pattern 3, n = 6). The primers
RT astA_1 and RT astA_2 matched the region between the 798th and 787th nucleotides, covering both the transcripts of astA or orf1 (Fig. 1). After qRT-PCR, each pattern showed distinct profiles of transcripts when compared to the prototype EAEC 042 (Table 1). Both pattern 1 strains showed a decreased level of transcripts. All but one strains of pattern 2 showed no differences in the transcripts level when compared to the prototype. Pattern 3 was more diverse, presenting strains with increased, decreased or similar levels of transcription.

DISCUSSION

Some putative virulence factors are shared between extraintestinal and diarrheagenic Escherichia coli. EAST1 is supposed to play a role in enteric diseases caused by certain diarrheagenic E. coli pathotypes. However, because the gene that encodes this protein is also detected in extraintestinal E. coli, we aimed to characterise it in isolates that are not associated with diarrhea.

The same DNA sequence that encodes EAST1 also encodes a putative transposase (ORF1) in a different frame (McVeigh et al. 2000). Strains presenting pattern 1 (astA pseudogene) encoded a truncated EAST1. On the other hand, the predicted transposase detected with pattern 1 is identical to a functional protein reported previously (McVeigh et al. 2000) and was not truncated. This indicates that the DNA sequence in these strains encodes a functional transposase and a non-functional EAST-1 protein. The underregulation detected in the transcripts within this pattern suggests that the role of this transposase is to transport copies of astA, and when astA is non-functional, the transposase production is downregulated.

Pattern 2 and pattern 3 did not show any exclusive transcriptional behaviour. Isolates presenting pattern 3 were normalregulated, upregulated or even downregulated. The upregulation detected in some strains may be due to bacterial regulators and/or to the fact that the gene astA can be found in multiple copies in chromosomes and plasmids (Yamamoto and Echeverria 1996; McVeigh et al. 2000). The downregulation may be due to regulators. Because isolates with distinct phylogenetic backgrounds possessed pattern 3, it is possible that some isolates harbour traits that replace the role of astA. Therefore, the isolates could upregulate these traits and downregulate astA.

The polymorphism presented in the present work indicates that the DNA sequences that encode ORF1 and astA are under intense selection. The region upstream astA was conserved among the extraintestinal strains presented in this work and also EAEC042 and EAEC 17-2. However, when compared with the sequences present in ETEC strains E68 and Abbostown, there were six point mutations (Yamamoto and Nakazawa 1997). Most of those mutations were neutral, but, considering the frame that encodes ORF1, there was a nucleotide substitution (GAT→AT) which yielded an amino acid with distinct chemical properties (Asp→Asn). The sequences downstream astA was also a target for mutations, but when comparing the extraintestinal E. coli strains with EAEC prototype strains, ETEC and DAEC, there was only synonymous mutations. The presence of extensive neutral substitutions suggests that the product of ORF1 (putative transposase) is persistent in E. coli. Since the transposase carries the gene astA, this gene might have a role that goes beyond virulence, particularly in extraintestinal isolates.

When comparing the central region which encodes astA, there were three substitutions. They were all neutral regarding the frame that encodes ORF1. However, for the frame that encodes astA, one mutation was silent, another produced a distinct amino acid and the third drove to a stop codon in the middle of the sequence. The stop codon mutation was only found in extraintestinal strains, while the other two were found in extraintestinal and also the EAEC prototypes. This stop codon was tested before in a previous work and proved to produce a non-toxic strain (McVeigh et al. 2000). It seems reasonable to assume that nonsense mutations generating non-toxic products are disfavoured to be maintained in strains whose role is to cause enteric diseases. The detection of astA genes and pseudogenes in extraintestinal isolates suggests that astA may have a role in the persistence of these isolates, but not as crucial as in diarrheagenic isolates.

The results show that the entire astA/ORF1 sequence has undergone selection. However, over the frame encoding ORF1, the selection was much more neutral than over the frame encoding astA. Apparently, the transposase function is well established while astA has evolved, perhaps to refine its function.

In conclusion, astA sequence patterns found in extraintestinal E. coli is similar to sequences found in diarrheagenic E. coli, including overlapping transposase encoding genes. However, extraintestinal E. coli may also present an astA pseudogene. Each sequence pattern was not associated with distinct levels of transcription, except with the pattern related to the pseudogene.

FUNDING

RPM (2012/05073-3) and TCCJ (2013/09167-5) are fellows from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). JLL was funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (PFPD/CAPES). WDS has a scientific fellowship from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). This work was funded by FAPESP (2012/04931-6).

Conflict of interest. None declared.

REFERENCES


Maluta RP, Logue CM, Casas MRT et al. Overlapped sequence types (STs) and serogroups of avian pathogenic (APEC) and human extra-intestinal pathogenic (ExPEC) Escherichia coli isolated in Brazil. PLoS One 2014;9:e105016.


