



# Antimicrobial resistance and gene regulation in Enteroaggregative Escherichia coli from Egyptian children with diarrhoea: Similarities and differences

Radwa Abdelwahab, Muhammad Yasir, Rita E. Godfrey, Gabrielle S. Christie, Sarah J. Element, Faye Saville, Ehsan A. Hassan, Entsar H. Ahmed, Nagla H. Abu-Faddan, Enas A. Daef, Stephen J. W. Busby & Douglas F. Browning

To cite this article: Radwa Abdelwahab, Muhammad Yasir, Rita E. Godfrey, Gabrielle S. Christie, Sarah J. Element, Faye Saville, Ehsan A. Hassan, Entsar H. Ahmed, Nagla H. Abu-Faddan, Enas A. Daef, Stephen J. W. Busby & Douglas F. Browning (2021) Antimicrobial resistance and gene regulation in Enteroaggregative Escherichia coli from Egyptian children with diarrhoea: Similarities and differences, *Virulence*, 12:1, 57-74, DOI: [10.1080/21505594.2020.1859852](https://doi.org/10.1080/21505594.2020.1859852)

To link to this article: <https://doi.org/10.1080/21505594.2020.1859852>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 29 Dec 2020.



[Submit your article to this journal](#)



Article views: 2129



[View related articles](#)

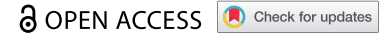


[View Crossmark data](#)




Citing articles: 5 [View citing articles](#)

RESEARCH PAPER



## Antimicrobial resistance and gene regulation in Enteroaggregative *Escherichia coli* from Egyptian children with diarrhoea: Similarities and differences

Radwa Abdelwahab<sup>a,b</sup>, Muhammad Yasir<sup>a,c</sup>, Rita E. Godfrey<sup>a</sup>, Gabrielle S. Christie<sup>a</sup>, Sarah J. Element<sup>a</sup>, Faye Saville<sup>a</sup>, Ehsan A. Hassan<sup>b</sup>, Entsar H. Ahmed<sup>b</sup>, Nagla H. Abu-Faddan<sup>b</sup>, Enas A. Daef<sup>b</sup>, Stephen J. W. Busby<sup>a</sup>, and Douglas F. Browning <sup>a</sup>

<sup>a</sup>Institute of Microbiology and Infection, School of Biosciences, University of Birmingham, Birmingham, UK; <sup>b</sup>Faculty of Medicine, Assiut University, Assiut, Egypt; <sup>c</sup>Quadram Institute Bioscience, Norwich Research Park, Norwich, UK

### ABSTRACT

Enteroaggregative *Escherichia coli* (EAEC) is a common diarrhoeagenic human pathogen, isolated from patients in both developing and industrialized countries, that is becoming increasingly resistant to many frontline antibiotics. In this study, we screened 50 *E. coli* strains from children presenting with diarrhea at the outpatients clinic of Assiut University Children's Hospital, Egypt. We show that all of these isolates were resistant to multiple classes of antibiotics and identified two as being typical EAEC strains. Using whole genome sequencing, we determined that both isolates carried, amongst others, *bla*<sub>CTX-M</sub> and *bla*<sub>TEM</sub> antibiotic resistance genes, as well as many classical EAEC virulence determinants, including the transcriptional regulator, AggR. We demonstrate that the expression of these virulence determinants is dependent on AggR, including *aar*, which encodes for a repressor of AggR, Aar. Since biofilm formation is the hallmark of EAEC infection, we examined the effect of Aar overexpression on both biofilm formation and AggR-dependent gene expression. We show that whilst Aar has a minimal effect on AggR-dependent transcription it is able to completely disrupt biofilm formation, suggesting that Aar affects these two processes differently. Taken together, our results suggest a model for the induction of virulence gene expression in EAEC that may explain the ubiquity of EAEC in both sick and healthy individuals.

### ARTICLE HISTORY

Received 28 August 2020  
Revised 24 November 2020  
Accepted 1 December 2020

### KEYWORDS

EAEC; antibiotic resistance; virulence; bacterial gene regulation; genome sequencing


## Introduction

Diarrhoeagenic *Escherichia coli* strains are important human pathogens, which cause considerable morbidity and mortality around the globe, particularly amongst infants and children in developing countries. These pathogens are classified into different pathotypes, which include enteroaggregative *E. coli* (EAEC), enterohaemorrhagic *E. coli* (EHEC), enteroinvasive *E. coli* (EIEC), enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC) and diffuse adhering *E. coli* (DAEC), based on their disease characteristics and specific adherence patterns [1]. Enteroaggregative *Escherichia coli* (EAEC) is a commonly isolated human pathogen that is responsible for causing mucoid diarrhea in patients from both industrialized and developing countries [2–5]. EAEC has been shown to elicit travelers' diarrhea, pediatric diarrhea, impairment of pediatric growth and cognition, and even extra-intestinal infections, such as urinary tract infections and septicemia [4,6–11]. In addition, EAEC strains

have been the cause of a number of serious diarrhea outbreaks throughout the world, with the food-borne outbreak caused by the Shiga-toxin-producing EAEC O104:H4 strain in Germany, resulting in 54 deaths [12–16].

EAEC pathogenesis proceeds by the colonization of the human intestinal mucosa followed by the secretion of toxins, such as plasmid-encoded toxin (Pet) and enteroaggregative heat-stable toxin (EAST-1). This, coupled with the resultant inflammation, is thought to lead to diarrhea and disease [17,18]. Typical EAEC strains carry the plasmid-encoded AggR transcription regulator protein, a member of the AraC-XylS family of transcription factors [19,20]. AggR activates the expression of many genes thought to be required for pathogenesis, for example the attachment adherence fimbriae (AAF) required for colonization, the anti-aggregation protein dispersin (Aap), and its dedicated type I secretion system (T1SS) [21–24]. Additionally, expression of AggR is up-regulated by AggR itself [25] and its activity is reported to be down-regulated by the Aar

**CONTACT** Stephen J. W. Busby  [s.j.w.busby@bham.ac.uk](mailto:s.j.w.busby@bham.ac.uk); Douglas F. Browning  [D.F.Browning@bham.ac.uk](mailto:D.F.Browning@bham.ac.uk)

 Supplemental data for this article can be accessed [here](#).

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

protein (AggR-activated regulator), whose expression is also induced by AggR [24,26,27]. Thus, there is a complex interplay between AggR and Aar to control EAEC virulence [26,28].

Much of what we understand about EAEC virulence comes from research with archetypal EAEC strains, such as EAEC strains 042 and 17–2 [29]. However, it is becoming clear that, in the environment, EAEC strains are extremely heterogeneous and many virulence determinants are not present in all strains [5,16]. For example, both EAEC 042 and 17–2 possess large virulence plasmids (*i.e.* pAA2 and pAA, respectively) and carry EAEC-associated genes encoding AggR, Aap and EAST-1 (*astA*). However, EAEC 042 possesses AAF Type II fimbriae (*aafDA* and *aafB-aafCB*), whilst EAEC 17–2 expresses AAF Type I fimbriae (*aggDCBA*) and lacks the Pet cytotoxin [5,19,21,30]. Like many enteric pathogens, EAEC strains are increasingly resistant to many clinically used antibiotics [31–33] and, as EAEC is often isolated from asymptomatic individuals [5,16], there is a need to understand how EAEC is tolerated by their human hosts but causes disease in certain individuals.

In Egypt, EAEC is increasingly recognized as an emerging enteric pathogen, showing high levels of antibiotic resistance. However, in-depth molecular characterization of Egyptian EAEC isolates has been limited to only a few studies [34–38]. In order to examine both the prevalence of antibiotic resistance and the occurrence of EAEC, we screened *E. coli* strains from infants and children with diarrhea from Assiut, Egypt. Alarmingly, most strains were resistant to many classes of antibiotics. Amongst these, we found two apparently typical EAEC strains that we characterized, using whole genome sequencing, identifying their virulence determinants, and investigating their regulation by AggR. We show that feed-forward activation of AggR expression by AggR is conserved, but that Aar-mediated repression is not. Our results lead to a model for the triggering of virulence in typical EAEC strains.

## Materials and methods

### Bacterial strains, plasmids, primers and growth conditions

The bacterial strains, plasmids and promoter fragments used in this study are listed in Table S1. The oligodeoxynucleotide primers used to amplify and mutate the various DNA fragments are listed in Table S2. Standard procedures for PCR, cloning and DNA manipulation were used throughout [39]. All promoter DNA fragments used in this study are flanked by EcoRI

and HindIII and the DNA sequence of each fragment is numbered from the base adjacent to the HindIII site. Base substitutions are defined by the position of the base altered and the substituted base introduced. Cells were routinely grown in Lysogeny Broth (LB medium) at 37°C with shaking. To measure promoter activities, fragments were cloned into the low copy number broad host range *lac* expression vectors, pRW50 [40] and pRW224 [41] and maintained with 15 µg ml<sup>-1</sup> tetracycline. To examine the effect of *aggR* and *aar* expression, cells were transformed with various pBAD derivatives (Table S1), which were maintained in cells with 100 µg ml<sup>-1</sup> ampicillin. AggR and/or Aar expression, using pBAD vectors, was induced by the addition of 0.2% w/v arabinose in the medium, where appropriate [22].

### Isolation and characterization of Egyptian *E. coli* strains

In total, 113 stool samples were collected from infants and children, whose age ranged between 2 months and 5 years, presenting with diarrhea to the outpatients clinic of Assiut University Children's Hospital in 2016. Ethical approval was granted by the Medical School Ethical Review Board before sample collection proceeded. Cases enrolled in this study had diarrhea, characterized by frequent watery stools (>3 times/day), with or without blood or mucus. Children and infants with severe protein energy malnutrition (either Marasmus or Kwashiorkor), or who had received antibiotics within the last 72 hours, were excluded from the study. Only one stool specimen from each case was tested, with isolation and identification of *E. coli* carried out at the Medical Research Center, Faculty of Medicine, Assiut University. All *E. coli* strains were tested for susceptibility to a range of antimicrobial agents, using the Kirby-Bauer disc diffusion method [42], which was interpreted according to the CLSI 2014 [43]. The antimicrobial discs (Hi-Media, India) contained the following antibiotics: imipenem (10 µg), meropenem (10 µg), trimethoprim/sulfamethoxazole (5 µg), cefaclor (30 µg), ceftriaxone (30 µg), ampicillin (10 µg), ciprofloxacin (5 µg), oxytetracycline (30 µg), amoxicillin (25 µg), norfloxacin (10 µg), tobramycin (10 µg) and amikacin (30 µg). To identify EAEC-associated genes, isolates were screened for *aggR*, *aap* and the CVD432 marker sequence (which encodes part of the EAEC Aat T1SS), using PCR and the primers detailed in Table S2.

### Complete genome sequencing and analysis

Complete genome sequencing of EAEC strains E36 and E42 was carried out using Illumina sequencing by

Microbes NG (<https://microbesng.com/>). Plated cultures of each isolate were inoculated into a cryopreservative (Microbank™, Pro-Lab Diagnostics UK). 10 to 20 µl of this suspension were lysed with 120 µL of TE buffer containing lysozyme (final concentration 0.1 mg mL<sup>-1</sup>) and RNase A (ITW Reagents, Barcelona, Spain) (final concentration 0.1 mg mL<sup>-1</sup>), incubated for 25 min at 37°C. Proteinase K (VWR Chemicals, Ohio, USA) (final concentration 0.1 mg mL<sup>-1</sup>) and SDS (Sigma-Aldrich, Missouri, USA) (final concentration 0.5% v/v) were added and incubated for 5 min at 65°C. Genomic DNA was purified using an equal volume of SPRI beads and resuspended in EB buffer (Qiagen, Germany). DNA was quantified with the Quant-iT dsDNA HS kit (ThermoFisher Scientific) assay in an Eppendorf AF2200 plate reader (Eppendorf UK Ltd, UK). Genomic DNA libraries were prepared using the Nextera XT Library Prep Kit (Illumina, San Diego, USA) following the manufacturer's protocol with the following modifications: 2 ng of DNA were used as input, and PCR elongation time was increased to 1 min from 30 s. DNA quantification and library preparation were carried out on a Hamilton Microlab STAR automated liquid handling system (Hamilton Bonaduz AG, Switzerland). Pooled libraries were quantified using the Kapa Biosystems Library Quantification Kit for Illumina on a Roche light cycler 96 qPCR machine. Libraries were sequenced with the Illumina HiSeq using a 250bp paired end protocol.

EAEC strain 17-2 was sequenced using the enhanced sequencing option from MicrobesNG, which utilizes both the Illumina and Oxford Nanopore Technologies (ONT). A broth culture was pelleted out and the pellet was resuspended in the cryopreservative of a Microbank™ (Pro-Lab Diagnostics UK, United Kingdom) tube and stored in the tube. Approximately  $2 \times 10^9$  cells were used for high molecular weight DNA extraction using Nanobind CCB Big DNA Kit (Circulomics, Maryland, USA). DNA was quantified with the Qubit dsDNA HS assay in a Qubit 3.0 (Invitrogen) Eppendorf UK Ltd, UK). Long read genomic DNA libraries were prepared with the Oxford Nanopore SQK-LSK109 kit with Native Barcoding EXP-NBD104/114 (ONT, UK), using 400–500 ng of HMW DNA. Twelve to twenty-four barcoded samples were pooled together into a single sequencing library and loaded on a FLO-MIN106 (R.9.4 or R.9.4.1) flow cell in a GridION (ONT, UK). Illumina reads were adapter trimmed using Trimmomatic 0.30 with a sliding window quality cutoff of Q15 [44]. Genome assembly was performed using Unicycler v0.4.0 [45] and contigs were annotated using Prokka 1.11 [46]. This Whole Genome Shotgun project

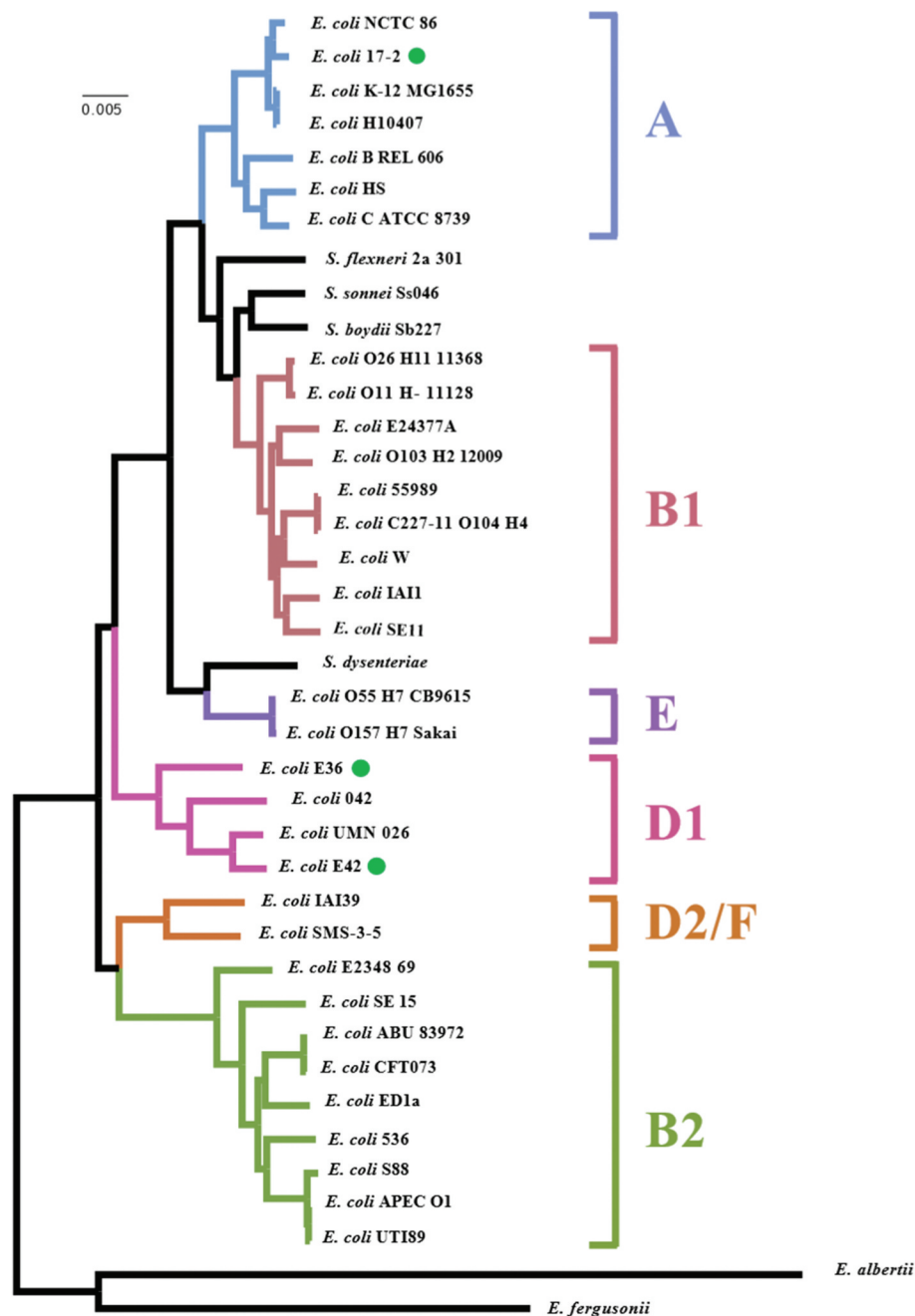
has been deposited at DDBJ/ENA/GenBank with the sequence data for EAEC strains E36, E42 and 17-2 under the accession numbers JACEFX000000000, JACEFW000000000 and JACEFV000000000, respectively.

### **Bioinformatic analysis of genome sequences**

Draft genomes were visualized using Artemis [47], comparisons between EAEC genomes were examined using the CGView Server ([http://stothard.afns.ualberta.ca/cgview\\_server/](http://stothard.afns.ualberta.ca/cgview_server/)) [48], the Basic Local Alignment Search Tool (BLAST) at NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) and the Artemis Comparison Tool (ACT) [49]. Representations of genome organization were drawn using the CGView Server [48], ACT [49] and DNAPlotter [50]. *E. coli* sequence types were determined using MLST 2.0 [51], bacterial serotyping was determined using SerotypeFinder 2.0 [52], plasmid replicons were detected using PlasmidFinder 2.1 [53], antibiotic resistance gene analysis used ResFinder 3.2 [54], and virulence gene analysis was performed using VirulenceFinder 2.0 [55] with the online software from the Center for Genomic Epidemiology (<http://www.genomicepidemiology.org/>). Insertion sequences were identified using ISfinder (<https://www-is.biotoul.fr/blast/resultat.php>) [56]. The phylogenetic analysis of EAEC strains E36 and E42 was carried out by recreating the phylogenetic tree from Dunne *et al.* [57], using the 35 genomes of *Escherichia* and *Shigella* listed in (Figure 1), the three EAEC genomes generated in this study, and EAEC O104:H4 strain C227-11 (accession number AFST000000000) [58]. Automated annotation was performed using Prokka v1.12 [46]. A pangenome analysis was performed using Roary v3.13.0 [59] to determine the sizes of the core genomes and pangenomes. A whole-genome phylogeny was reconstructed from the core genome alignment using RAxML 8.2.4 [60].

### **Promoter fragment and plasmid construction**

The promoter fragments *aggR90*, *aar100*, *aatP100* and *aatP98* were amplified by PCR using the primer pairs listed in Table S2, with EAEC 042, EAEC 17-2, EAEC E36 or EAEC E42 genomic DNA as template. All DNA fragments are flanked by EcoRI and HindIII sites to facilitate cloning into pRW50 and pRW224 to generate *lacZ* transcriptional fusions [40,41], and sequences are numbered from the HindIII site. Point mutations were introduced into fragments using megaprimer PCR [61] and constructs were verified by Sanger DNA sequencing.



**Figure 1.** Phylogenetic analysis of EAEC strains E36 and E42. The figure shows a phylogenetic tree of *E. coli* strains highlighting the position of EAEC strains E36, E42 and 17–2 (green dots). The phylogenetic tree was reconstructed from Dunne *et al.* [57], using RAxML (version 8.2.4) based on maximum likelihood analysis. Phylogenetic groups and labeling is similar to the reference phylogenetic tree in [57] and EAEC O104:H4 strain C227-11 (accession number AFST00000000) [58] has also been included.

To generate the various pBAD30/*aar* plasmids used in this study (Table S1), the *aar* gene was amplified from EAEC 042 and EAEC E36 DNA by PCR using primers *aar* XbaI up, *aar* E36 XbaI up or *aar*\* XbaI up with primers *aar* SphI down or *aar* E36 SphI down, to generate each *aar* or *aar*\* PCR products (Tables S1 and S2). Note that *aar*\* derivatives carry the strong RBS from the pET20b expression vector (Novagen).

Purified PCR products were restricted with XbaI and SphI and then cloned into pBAD30, cut with the same enzymes, thus, placing the expression of *aar* derivatives under the control of the arabinose inducible *paraBAD* promoter [62]. To generate the pBAD/*aggR/aar* series of vectors (Table S1), purified *aar* and *aar*\* PCR products were cloned downstream of *aggR* in pBAD/*aggR*, again using XbaI and SphI, to generate synthetic *aggR*-

*aar* operons, which allow the concurrent expression of AggR and Aar, by induction with 0.2% w/v arabinose.

### Assays of promoter activity

To assay the expression from promoter derivatives cloned into the *lac* expression vectors pRW50 and pRW224, *E. coli* K-12 BW25113  $\Delta lac$  strain was transformed with each construct and  $\beta$ -galactosidase activity was measured, as described in our previous work [27,63]. AggR was expressed from pBAD/*aggR*, which carries EAEC 042 *aggR* cloned downstream of the arabinose inducible promoter, *paraBAD*[22]. Cells were grown in LB medium at 37°C with shaking to mid-logarithmic phase ( $OD_{650} = 0.4$ – $0.6$ ) and 0.2% w/v arabinose was included in the medium to induce AggR expression, where appropriate.  $\beta$ -galactosidase activities are expressed as nmol of ONPG hydrolyzed  $\text{min}^{-1} \text{mg}^{-1}$  dry cell mass and each activity is the average of three independent determinations.

### Quantitative biofilm formation assays

Quantitative biofilm formation assays were carried out as in our previous work [5]. Overnight EAEC cultures, grown in LB medium, were sub-cultured (1 in 100) into 5 ml of DMEM (Dulbecco's Modified Eagle Medium) high glucose (Sigma) and incubated at 37°C for 1 hour with shaking. 150  $\mu\text{l}$  of each culture was pipetted into a microtiter plate in triplicate, which was sealed with a Breath Easy gas permeable membrane (Sigma), and then incubated overnight statically at 37°C. After ~16–17 hrs the spent media was removed and 150  $\mu\text{l}$  of 0.1% (w/v) crystal violet was added to each well and left at 4°C for 30 minutes. The crystal violet solution was removed, the plate washed thoroughly with water, excess liquid removed and 150  $\mu\text{l}$  of ethanol/acetone solution (80 ml ethanol and 20 ml acetone) was added. The plate was left on a shaker for 30 minutes at room temperature and the absorbance was measured at 595 nm by a Labsystems Multiskan MS plate reader (Thermo Fisher Scientific Inc).

## Results and discussion

### Isolation and characterization of *E. coli* strains from Egyptian children with diarrhea

In total, 50 *E. coli* strains were isolated from infants and children, presenting with diarrhea at the outpatient clinic of Assiut University Children's Hospital, and their susceptibility against various antimicrobials tested. Results in Table 1 show that all were resistant to the

cephalosporin antibiotics, cefaclor (100%) and ceftriaxone (100%), and the vast majority of isolates were resistant to trimethoprim/sulfamethoxazole (94%) and ampicillin (98%). Furthermore, four isolates (E16, E28, E35 and E36) were resistant to all tested antibiotics (Table S3), highlighting the multidrug resistance phenotype of many of the *E. coli* strains isolated during this study.

To determine if any of these isolates were EAEC in character, strains were screened for the EAEC-associated genes, *aggR* and *aap*, as well as the CVD432 marker sequence, which encodes part of the EAEC Aat T1SS [10]. Two strains, E36 and E42, possessed *aggR*, *aap* and CVD432, identifying them as typical EAEC isolates. To verify this, and to determine what antibiotic resistance genes and virulence determinants they carried, the genome of each of these strains was sequenced, together with the genome of the prototypical EAEC strain, 17–2, as a reference (Table S1). Analysis indicated that the genomes of strains E36 and E42 were similar in size, with E42 containing slightly more coding sequences (*i.e.* 5046 vs 4912) (Table 2). Both strains were found to belong to *E. coli* phylogenetic group D, like EAEC 042, with E36 belonging to sequence type ST38 and E42 to ST1380 (Table 2; Figure 1). EAEC 17–2, which possessed a slightly smaller genome, was *E. coli* phylogenetic group A and sequence type ST10 (Table 2; Figure 1) [64]. For each strain, a number of plasmid replicons were detected, consistent with each strain possessing a number of plasmids [53]. Due to the draft nature of the E36 and E42 genome assemblies, each plasmid replicon was found on separate contigs and, therefore, it is not immediately clear how many plasmids each isolate possesses. However, analysis of EAEC 17–2 (Table 2), which possess a single large virulence plasmid (pAA) [65], revealed

**Table 1.** Antimicrobial susceptibility profile of 50 *E. coli* strains isolated from children with diarrhea from Assiut Children's Hospital, Egypt.

| Antibiotic                    | Resistant<br>n (%) | Sensitive<br>n (%) | Intermediate <sup>a</sup><br>n (%) |
|-------------------------------|--------------------|--------------------|------------------------------------|
| Imipenem                      | 16 (32)            | 30 (60)            | 4 (8)                              |
| Meropenem                     | 14 (28)            | 27 (54)            | 9 (18)                             |
| Trimethoprim/sulfamethoxazole | 47 (94)            | 2 (4)              | 1 (2)                              |
| Cefaclor                      | 50 (100)           | 0 (0)              | 0 (0)                              |
| Ceftriaxone                   | 50 (100)           | 0 (0)              | 0 (0)                              |
| Ampicillin                    | 49 (98)            | 1 (2)              | 0 (0)                              |
| Ciprofloxacin                 | 24 (48)            | 13 (26)            | 13 (26)                            |
| Oxytetracycline               | 29 (58)            | 14 (28)            | 7 (14)                             |
| Amoxicillin                   | 27 (54)            | 12(24)             | 11 (22)                            |
| Norfloxacin                   | 22 (44)            | 21(42)             | 7 (14)                             |
| Tobramycin                    | 34 (68)            | 1 (2)              | 15 (30)                            |
| Amikacin                      | 22 (44)            | 13 (26)            | 15 (30)                            |

<sup>a</sup>Intermediate resistance refers to strains above the point of antibiotic susceptibility but below the resistant breakpoint as defined by the [43].

**Table 2.** Genome analysis of EAEC isolates E36, E42, 17–2, 042 and C227-11.

|  | EAEC E36   | EAEC E42  | EAEC 17–2  | EAEC 042   | EAEC C227-11  |
|--|--|---|--|--|---|
| Genome size                              | 5,410,238 bp   | 5,290,086 bp  | 5,185,323 bp   | 5,355,323 bp   | 5542971bp   |
| G/C %                                    | 50.3%  | 50.47%  | 50.63%   | 50.56% <sup>h</sup>  | 50.71% <sup>h</sup>   |
| Number of contigs                        | 150  | 149   | 9  | N/A  | 37  |
| Genes (CDS)                              | 5046   | 4912  | 4924   | 4921   | 5431  |
| Phylotype                                | D  | D   | A  | D  | B1  |
| Sequence Type <sup>a</sup>               | ST38   | ST1380  | ST10   | ST414  | ST678   |
| Serotype <sup>b</sup>                    | O?:H30   | O17:H18   | O3:H2  | O44:H18  | O104:H4   |
| Plasmid replicons <sup>c</sup>           | IncFIB, IncFII, IncQ1  | IncFIB, IncFII, IncI1-I (γ)   | IncFIB, IncFII   | IncFIC(FII)  | IncFIB, IncFII, IncI1-I(γ), IncQ1   |
| Antibiotic resistance genes <sup>d</sup> | <i>aadA1</i> , <i>aph(3'')-Ib</i> , <i>aph(3')-Ia</i> , <i>aph(6)-Id</i> , <i>bla<sub>CTX-M-14b</sub></i> , <i>bla<sub>TEM-1B</sub></i> , <i>dfrA1</i> , <i>mdfA</i> , <i>sul2</i> , <i>tetA</i>   | <i>bla<sub>CTX-M-15</sub></i> , <i>bla<sub>TEM-1B</sub></i> , <i>mdfA</i> , <i>gyrA</i> pS83A   | <i>aadA1</i> , <i>dfrA1</i> , <i>mdfA</i> , <i>sul2</i>  | <i>aadA1</i> , <i>catA1</i> , <i>mdfA</i> , <i>sul1</i> , <i>tetA</i>  | <i>aph(3'')-Ib</i> , <i>aph(6)-Id</i> , <i>bla<sub>CTX-M-15</sub></i> , <i>bla<sub>TEM-1B</sub></i> , <i>dfrA7</i> , <i>sul1</i> , <i>sul2</i> , <i>mdfA</i> , <i>tetA</i> , <i>gyrA</i> pS83A  |
| Virulence genes <sup>e,f</sup>           | <b><i>orf3</i></b> , <b><i>orf4</i></b> , <b><i>aap</i></b> , <b><i>aar</i></b> , <b><i>aatA</i></b> , <b><i>agg3B</i></b> , <b><i>agg3C</i></b> , <b><i>agg3D</i></b> , <b><i>agg5A</i></b> , <b><i>aggR</i></b> , <b><i>capU</i></b> , <i>air</i> , <i>eilA</i> , <i>gadB</i> , <i>chuA</i> , <i>fyuA</i> , <i>irp2</i> , <i>iss</i> | <b><i>orf3</i></b> , <b><i>orf4</i></b> , <b><i>aaiC</i></b> , <b><i>aap</i></b> , <b><i>aatA</i></b> , <b><i>agg4A</i></b> , <b><i>agg4C</i></b> , <b><i>agg4D</i></b> , <b><i>aggR</i></b> , <b><i>capU</i></b> , <i>air</i> , <i>eilA</i> , <i>gadA</i> , <i>gadB</i> , <i>chuA</i> , <i>iss</i> , <i>lpfA</i> , <i>sepA</i> | <b><i>orf3</i></b> , <b><i>orf4</i></b> , <b><i>aaiC</i></b> , <b><i>aap</i></b> , <b><i>aar</i></b> , <b><i>aatA</i></b> , <b><i>aggA</i></b> , <b><i>aggB</i></b> , <b><i>aggC</i></b> , <b><i>aggD</i></b> , <b><i>aggR</i></b> , <b><i>capU</i></b> , <i>astA</i> , <i>gadA</i> , <i>gadB</i> , <i>iha</i> , <i>iss</i> , <i>fyuA</i> , <i>irp2</i> , <i>iucC</i> , <i>iutA</i> , <i>papA</i> , <i>papC</i> , <i>sat</i> | <b><i>orf3</i></b> , <b><i>orf4</i></b> , <b><i>aaiC</i></b> , <b><i>aap</i></b> , <b><i>aar</i></b> , <b><i>aatA</i></b> , <b><i>aafA</i></b> , <b><i>aafB</i></b> , <b><i>aafC</i></b> , <b><i>aafD</i></b> , <b><i>aggR</i></b> , <b><i>capU</i></b> , <i>astA</i> , <i>air</i> , <i>eilA</i> , <i>gadA</i> , <i>gadB</i> , <i>chuA</i> , <i>fyuA</i> , <i>irp2</i> , <i>lpfA</i> , <i>mchB</i> , <i>mchC</i> , <i>mchF</i> , <i>mcmA</i> , <i>pet</i> , <i>pic</i> | <b><i>orf3</i></b> , <b><i>orf4</i></b> , <b><i>aaiC</i></b> , <b><i>aap</i></b> , <b><i>aar</i></b> , <b><i>aatA</i></b> , <b><i>aggA</i></b> , <b><i>aggB</i></b> , <b><i>aggC</i></b> , <b><i>aggD</i></b> , <b><i>aggR</i></b> , <b><i>capU</i></b> , <i>gadA</i> , <i>gadB</i> , <i>iha</i> , <i>fyuA</i> <i>irp2</i> , <i>iucC</i> , <i>iutA</i> , <i>lpfA</i> , <i>mchB</i> , <i>mchC</i> , <i>mchF</i> , <i>neuC</i> , <i>pic</i> , <i>sepA</i> , <i>sigA</i> , <i>stx2A</i> , <i>stx2B</i> |

Software at the Center for Genomic Epidemiology (<http://www.genomicepidemiology.org/>) was used to identify: <sup>a</sup> the sequence type [51], <sup>b</sup> the serotype [52], <sup>c</sup> the number of plasmid replicons [53], <sup>d</sup> the various antibiotic resistance genes [54] and <sup>e</sup> the potential virulence genes [55] of each strain. <sup>f</sup>Genes shown to be regulated by AggR are shown bold [24,27]. <sup>h</sup> G/C content of chromosome is given.

similar IncFIB and IncFII replicons on the same large 81 kbp contig (Fig. S1) [53]. As both of these replicons often occur together on other IncF plasmids [53,66], it is likely that the genomes of E36 and E42 contain a similar large virulence plasmid, bearing both IncF replicons. Interestingly, comparison of plasmid pAA2 from EAEC 042 [67] with our pAA sequence from EAEC 17–2, indicated that these two plasmids were very different, with regions of homology confined mainly to EAEC virulence determinants (see below) and insertion sequences (Fig. S1). This is consistent with the proposition that a conserved plasmid backbone may not exist for pAA plasmids, and that EAEC virulence determinants have integrated into many different plasmids to produce the EAEC pathotype [64].

### Antibiotic resistance genes carried by EAEC strains E36 and E42

In the genome of EAEC E36, we detected various antibiotic resistance genes (Table 2), which can result in aminoglycoside resistance (*aadA1*, *aph(3'')-Ib*, *aph(3')-Ia* and *aph(6)-Id*), tetracycline resistance (*tetA*), β-lactam resistance (*bla<sub>CTX-M-14b</sub>* and *bla<sub>TEM-1B</sub>*), trimethoprim resistance (*dfrA1*), macrolide resistance (*mdfA*) and sulfonamide resistance (*sul2*). This is in line with the observed multidrug resistance phenotype of EAEC E36 (Table S3). For EAEC E42, we detected resistance genes for β-

lactams (*bla<sub>CTX-M-15</sub>* and *bla<sub>TEM-1B</sub>*), macrolides (*mdfA*) and nalidixic acid/ciprofloxacin (*gyrA* pS83A). Note that the extended spectrum β-lactamases (ESBL) type CTX-M and TEM found in both strains are often associated with EAEC isolates and the sequence types to which each strain belongs (*i.e.* ST38 and ST1380) [31–33,68].

Interestingly, in EAEC E36 the *sul2*, *aph(3'')-Ib* and *aph(6)-Id* resistance genes were found on the same 4996 bp contig as the IncQ1 plasmid replicon, indicating that they were likely plasmid encoded (Fig. S2). The organization of this replicon and resistance genes is identical to that of the broad-host range plasmid pRSF100 (Fig. S2) [53,69]. Due to the draft nature of the EAEC E36 genome sequence, it is unclear whether its ESBL genes are plasmid-borne, however, the *tetA*, *tetR* and *bla<sub>TEM-1B</sub>* genes are located on the same 3907 bp contig (Fig. S2), indicating that they would be co-inherited if plasmid encoded. In contrast, both ESBL genes from EAEC E42 (*i.e.* *bla<sub>CTX-M-15</sub>*, *bla<sub>TEM-1B</sub>*) were found on the same large contig as the IncI1-I(γ) plasmid replicon (Fig. S3), indicating that these genes are carried by the same plasmid. Indeed, IncI plasmids, isolated from animals and humans, have been shown to carry both *bla<sub>CTX-M-15</sub>* and *bla<sub>TEM-1</sub>* ESBLs, including the EAEC O104:H4 German outbreak strain C227-11 (Table 2) [58,70–72]. Furthermore, the entire sequence of this contig was 99% identical to a large resistance plasmid, pEC\_Bactec (92,970 bp) [72],

which was isolated in the U.K. from a horse, and carries a Tn3-like transposon which contains both ESBL genes (Fig. S3).

### Virulence genes carried by EAEC strains E36 and E42

Comparison of the E36 and E42 genome sequences with the EAEC 042 chromosome (Figure 2) and pAA2 plasmid (Figure 3) confirmed that both are typical EAEC strains (Table 2), possessing the AggR master regulator, which is very similar to AggR from EAEC strains 042 and 17-2 (Fig. S4) [19,21,67]. The EAEC E36 and 17-2 genomes also encode Aar, which can repress AggR mediated-activation [26,73] (Figure 3 and S4). However, Aar was absent from the genome of EAEC E42.

Both EAEC E36 and E42 possess characteristic EAEC AAF fimbriae genes, which are central to the attachment of EAEC strains to human epithelial cells [10], with E36 carrying AAF/V (*agg3DCB-agg5A*) and E42 AAF/IV (*agg4DCA*) type fimbriae (Table 2) [74,75]. Our analysis also confirmed that E36, E42 and 17-2 possess the anti-aggregation protein, dispersin (*aap*) [22], the Aat T1SS required for dispersin secretion (*aatA*) [23], *orf3* and *orf4* that encode proteins that are involved in biofilm formation [24] and *capU*, encoding a glycosyl transferase that forms an operon with *shf* and *virK* in EAEC 042 (Table 2; Figures 2 and 3) [76]. Like EAEC 042, EAEC E42 and 17-2 also carry the genes encoding the Aai type VI secretion system (T6SS) (*aaiC*) [77]. As all of these genes are induced by AggR in EAEC 042, it is likely that they also form part of the AggR regulon in EAEC E36, E42 and 17-2 [24,27]. Note that, in EAEC E36, all these genes (with the exception of *capU*) are located on two large contigs, which carried either the IncFIB plasmid replicon and/or plasmid conjugation genes (*i.e.* *tra* genes), indicating that they are plasmid-encoded, as is the case for typical EAEC strains [10]. As expected for strain EAEC 17-2, all AggR-regulated genes, with the exception of *aaiC* and *capU*, were found on the single large contig relating to the EAEC 17-2 virulence plasmid, pAA.

In addition to potential AggR-regulated virulence determinants, EAEC E36 and E42 both carry *air* (encoding the enteroaggregative immunoglobulin repeat protein), which is an accessory adhesin of EAEC 042 [78], *eilA*, which encodes the regulator of *air* expression [78], various glutamate decarboxylases genes (*gadA/gadB*) involved in acid resistance [79], and *iss/bor*, which encodes a lipoprotein involved in increased serum survival [80] (Table 2; Figure 2).

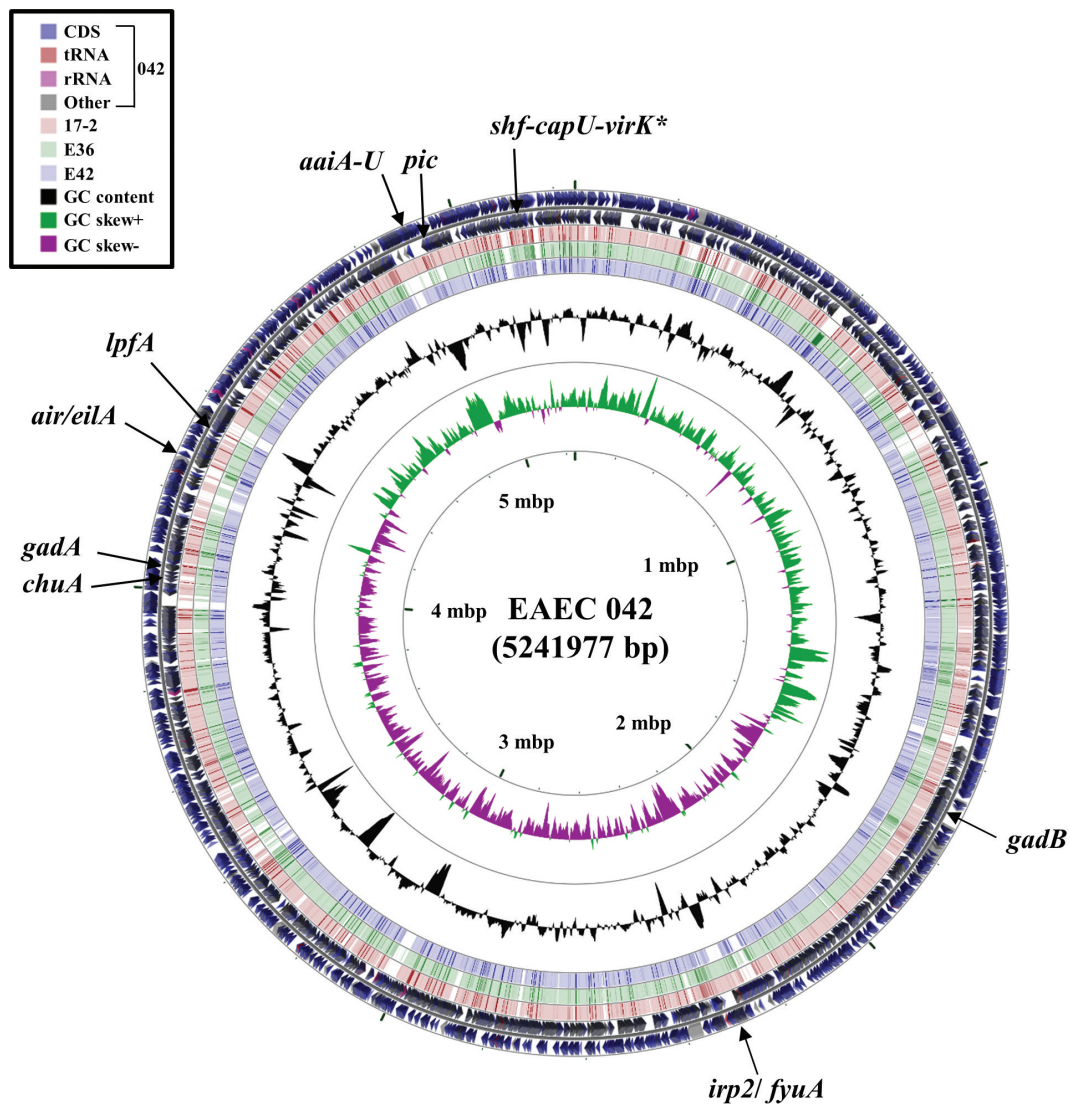
Being *E. coli* phylogenetic group D, like EAEC 042, both E36 and E42 carry the *chuA* heme-binding protein [67,81], whilst only E36 possesses the Yersiniabactin uptake system (*ipr2/fyuA*) found in EAEC 042 and 17-2 [67] (Table 2; Figure 2). In addition, EAEC E42 possesses the genes encoding the *lpfA* (long polar fimbriae) adhesin [82] and the *Shigella* extracellular protein A autotransporter toxin, *sepA* [83] (Table 2). Thus, it is clear that both EAEC E36 and E42 possess many genes associated with EAEC pathogenicity, as well as additional ones found in other pathogenic Enterobacteriaceae.

### Characterization of AggR-dependent promoters from different EAEC strains

Previously we predicted and characterized many AggR-dependent promoters from EAEC strain 042, identifying both the DNA site for AggR and the associated -10 promoter element [27]. Our studies showed that AggR binds to the consensus sequence 5'-WWWWWWWTATC-3' (where W = A/T) and functional sites are located 21 to 23 bp upstream of the -10 promoter hexamer element, which is recognized by the RNA polymerase  $\sigma$  subunit [27,84]. Comparison of the DNA sequences upstream of the transcript units encoding virulence genes in E36 (*aagR*, *aatP*, *aar*, *aap*, *orf3* and *agg3D*) and E42 (*aagR*, *aatP*, *aap*, *orf3*, *agg4D* and *aaiA*) (Table 2) identified similar promoters (Fig. S5), suggesting that all these genes formed part of the AggR regulon in these organisms.

To confirm our predictions, we focused on the *aatP* promoter region, as this promoter has not been previously characterized. During infection, EAEC produce the anti-aggregation protein dispersin (Aap) and its dedicated T1SS (AatPABCD), which aids in the presentation of AAF fimbriae for binding to human epithelial cells [22,23]. Thus, to examine the regulation of the *aatP* promoter, the DNA upstream of *aatP* was amplified by PCR from EAEC strains 042, E36 and E42, generating the *aatP*100 042, *aatP*100 E36 and *aatP*100 E42 promoter fragments, respectively (Table S1). Each fragment was cloned into the low copy number *lacZ* expression vector pRW50 [40] to generate a *lac* transcriptional fusion and plasmid constructs were transferred into the  $\Delta lac$  *E. coli* K-12 strain, BW25113. To investigate the effect of AggR, host cells also carried plasmid pBAD/*aggR*, which encodes AggR from EAEC 042 expressed from an arabinose-inducible promoter, or empty pBAD24 vector (Table S1) [22]. Cells were grown in LB medium to mid-logarithmic phase, either with or without AggR induction by arabinose, and  $\beta$ -galactosidase activities were determined as a proxy for promoter activity. Results illustrated in Figure 4 show



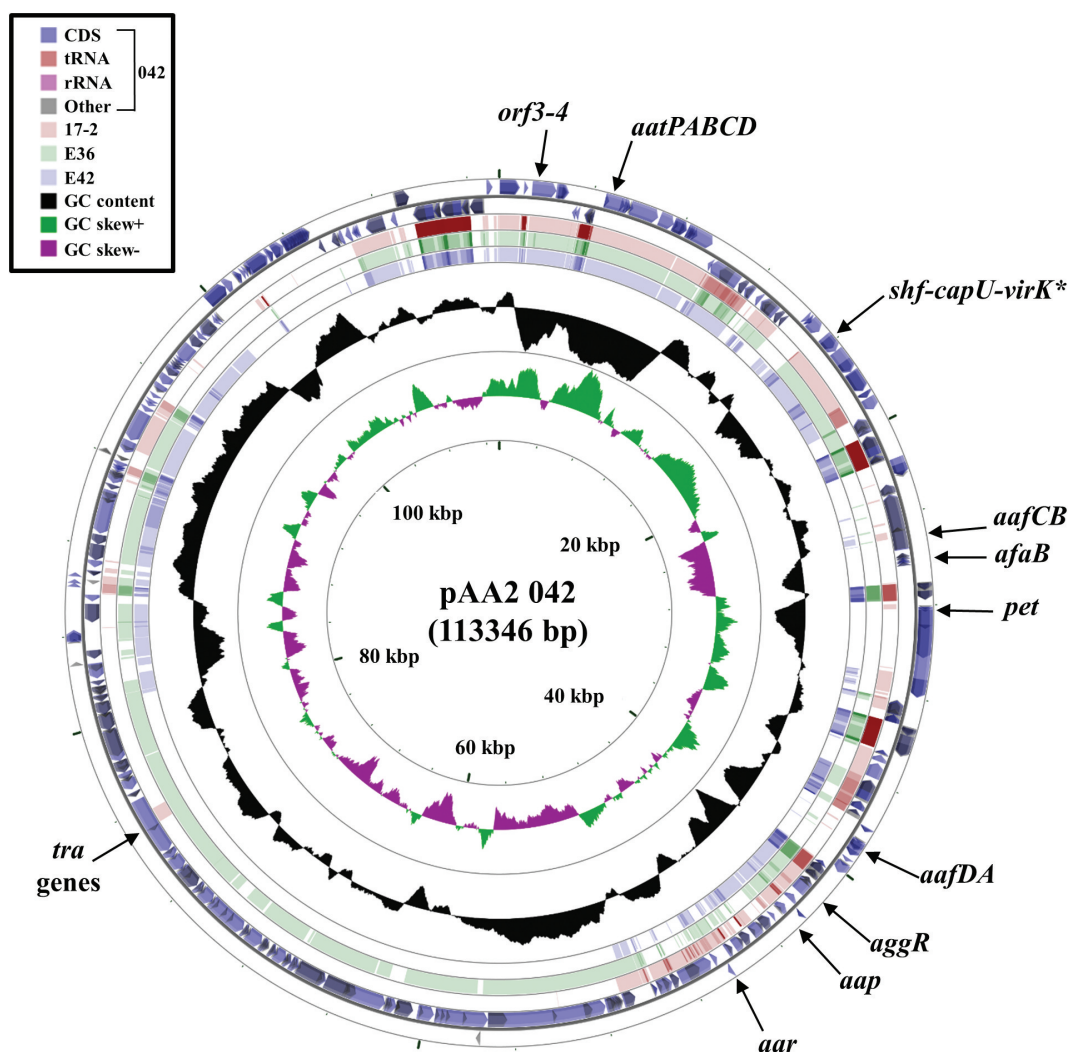


**Figure 2.** Comparison of the EAEC 042 chromosome with the genomes of EAEC strains 17–2, E36 and E42. The figure shows the comparison of the EAEC 042 chromosome with the genomes of EAEC 17–2, E36 and E42 using GCview ([http://stothard.afns.ualberta.ca/cgview\\_server/](http://stothard.afns.ualberta.ca/cgview_server/)) [48]. The outer two rings display the genes and features of the EAEC 042 chromosome (FN5554766.1) on both strands, with selected genes labeled [67]. The brown, green and blue rings illustrate the BLAST results when the genome sequences of EAEC 17–2, E36 and E42, respectively, are compared to the EAEC 042 genome, with shaded regions indicating synteny. The inner two rings display GC content (black) and GC skew (dark green and purple) for the EAEC 042 chromosome. The *shf-capU-virK* locus is starred as this region is found on both the EAEC 042 chromosome and pAA2 plasmid [67].

that expression from each *aatP100* construct was induced by AggR, confirming that the expression of the Aat T1SS in E36 and E42 is AggR-dependent. Note that, for both the E36 and E42 *aatP* promoters, there is a clear increase in AggR-independent activity (Figure 4), indicating that DNA sequence differences at these promoters affect expression (Fig. S5B). Furthermore, our assignments for the AggR-binding site and the –10 promoter element at the *aatP* promoter were confirmed by mutational analysis. Thus, point mutations, introduced into the predicted AggR-binding site and –10 element of the 042 *aatP* promoter

fragment greatly decreased AggR-dependent induction (Fig. S6).

Since AggR and Aar proteins are central to the regulation of EAEC virulence, we chose to examine the *aggR* and *aar* promoters in more detail. AggR expression in EAEC 042 is auto-activated [25], and alignment of the DNA upstream of *aggR* from EAEC E36, E42 and 17–2, with that from EAEC 042 [25], revealed polymorphisms in both the AggR-binding sites and the promoter –10 elements (Fig. S5). Hence, *aggR* promoter DNA from EAEC E36 and E42 was amplified using PCR to generate the *aggR90* E36 and

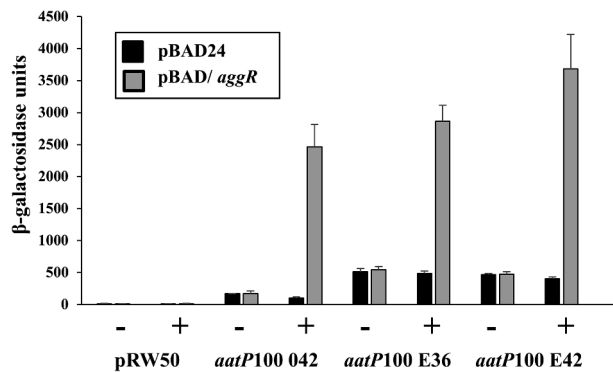


**Figure 3.** Comparison of the EAEC 042 plasmid pAA2 with the genomes of EAEC strains 17–2, E36 and E42. The figure shows the comparison of the EAEC 042 pAA2 plasmid with the genomes of EAEC 17–2, E36 and E42 using GCview ([http://stothard.afns.ualberta.ca/cgview\\_server/](http://stothard.afns.ualberta.ca/cgview_server/)) [48]. The outer two rings display the genes and features of pAA2 (FN554767.1) on both strands, with selected genes labeled [67]. The brown, green and blue rings illustrate BLAST results when the genome sequences of EAEC 17–2, E36 and E42, respectively, are compared to pAA2. The inner two rings display GC content (black) and GC skew (dark green and purple) for pAA. The *shf-capU-virK* locus is starred as this region is found on both the EAEC 042 chromosome and pAA2 [67].

*aggR90* E42 promoter fragments, which were cloned to generate plasmid-borne *lac* transcriptional fusions (Table S1). As controls, similar promoter fragments were generated from EAEC 042 and 17–2 (*i.e.* *aggR90* 042 and *aggR90* 17–2). Each construct was then transferred into BW25113 cells, carrying either pBAD/*aggR* or pBAD24.  $\beta$ -galactosidase activities were measured, as before and data in Figure 5 show that expression from each *aggR* promoter fragment was substantially increased by arabinose in cells containing pBAD/*aggR*, but not with pBAD24 alone. This indicates that each fragment carries an AggR-activated promoter. Note that the basal AggR-independent expression differs

from promoter to promoter, with the *aggR* promoter from EAEC 17–2 showing the highest background levels (Figure 5).

Although expression of the *aar* repressor protein has been examined by transcriptomic analysis, the *aar* promoter region had not been characterized [24,27] (Fig. S5). Therefore, the DNA upstream of *aar* was amplified from EAEC strains 042 and E36, to generate the *aar100* 042 and *aar100* E36 promoter fragments, which were then cloned to generate *lac* transcriptional fusions. The  $\beta$ -galactosidase activity of BW25113 cells, carrying these constructs, with pBAD/*aggR* or pBAD24, was then measured and data in Figure 6 show that



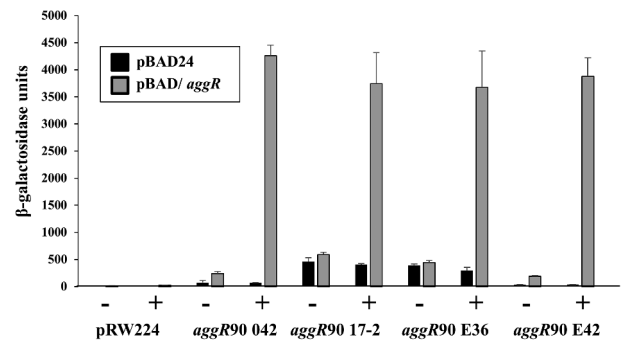
**Figure 4.** Activation of the *aatP* promoter from EAEC strains 042, E36 and E42. The figure illustrates measured  $\beta$ -galactosidase activities in *E. coli* K-12 BW25113  $\Delta lac$  cells, containing pRW50 carrying *aatP100* promoter fragments from EAEC strains 042, E36 and E42. Cells also carried either pBAD/*aggR* (gray bars) or pBAD24 (black bars), and were grown in LB medium with (+) or without (-) 0.2% arabinose.  $\beta$ -galactosidase activities are expressed as nmol of ONPG hydrolyzed min<sup>-1</sup> mg<sup>-1</sup> dry cell mass. Each activity is the average of three independent determinations and standard deviations are shown for all data points.

promoter activity from both *aar* promoters is increased by AggR, particularly for EAEC E36. Interestingly, for both *aar* constructs substantial expression is observed in the absence of AggR. To confirm our promoter predictions, point mutations were introduced into the AggR-binding site and -10 promoter element of the *aar100* 042 promoter fragment. Fig. S7 shows that these substitutions substantially decrease AggR-dependent and AggR-independent promoter activity, confirming our assignment of the key promoter elements controlling *aar* expression.

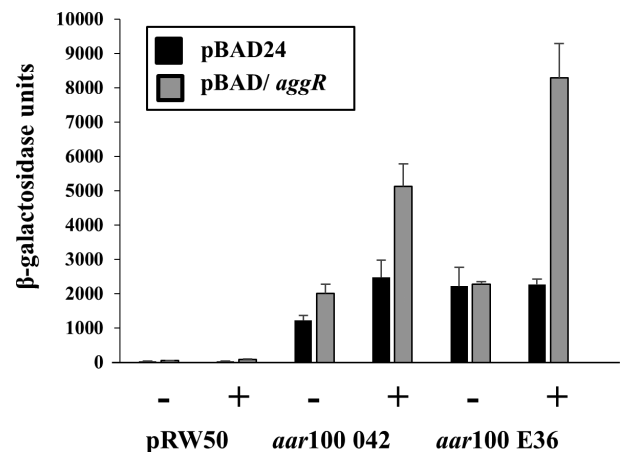
### The effect of Aar expression on EAEC biofilm formation

A characteristic of EAEC strains is their capacity to form biofilms on epithelial cells or abiotic surfaces [85,86]. Therefore, we investigated the ability of EAEC strains E36, E42, 042 and 17-2 to form biofilms on plastic microtitre plates when grown in DMEM high glucose medium, which induces AggR-mediated biofilm formation [5,24,27,85]. Results presented in Figure 7a show that EAEC E36 and E42, along with EAEC 17-2, produce biofilm, but to a lesser extent than EAEC 042. Such differences have been observed before for different EAEC isolates [5,85] and may reflect the ability of different AAF fimbrial types to adhere to abiotic surfaces.

Biofilm formation in EAEC 042 is dependent on AggR and the induction of the AggR-regulon (Figure



**Figure 5.** Activation of the *aggR* promoter from EAEC strains 042, 17-2, E36 and E42. The figure illustrates measured  $\beta$ -galactosidase activities in *E. coli* K-12 BW25113  $\Delta lac$  cells, containing pRW224 carrying *aggR90* promoter fragments from EAEC strains 042, 17-2, E36 and E42. Cells also carried either pBAD/*aggR* (gray bars) or pBAD24 (black bars) and were grown in LB medium in the presence (+) or absence (-) of 0.2% arabinose.  $\beta$ -galactosidase activities are expressed as nmol of ONPG hydrolyzed min<sup>-1</sup> mg<sup>-1</sup> dry cell mass. Each activity is the average of three independent determinations and standard deviations are shown for all data points.



**Figure 6.** Activity of the *aar* promoter from EAEC strains 042 and E36. The figure illustrates measured  $\beta$ -galactosidase activities in *E. coli* K-12 BW25113  $\Delta lac$  cells, containing pRW50 carrying *aar100* promoter fragments from EAEC strains 042 and E36. Cells also carried either pBAD/*aggR* (gray bars) or pBAD24 (black bars) and were grown in LB medium in the presence (+) or absence (-) of 0.2% arabinose.  $\beta$ -galactosidase activities are expressed as nmol of ONPG hydrolyzed min<sup>-1</sup> mg<sup>-1</sup> dry cell mass. Each activity is the average of three independent determinations and standard deviations are shown for all data points.

7b) [85,86]. Thus, to investigate the effect of Aar overexpression on biofilm formation, DNA fragments, carrying the *aar* coding sequence from EAEC strains 042 and E36, were amplified and cloned into the arabinose inducible vector pBAD30 (Table S1; Fig. S8) to generate

pBAD30/*aar* 042 and pBAD30/*aar* E36. Both constructs were transferred into EAEC 042 and the biofilm assays were repeated with arabinose added to induce Aar expression. Results in [Figure 7c](#) show that induction of E36 Aar decreased EAEC 042 biofilm production, whilst induction of 042 Aar expression had little or no effect. Comparison of the E36 and 042 *aar* sequences (Figs. S8A and B) identified the ribosome binding site (RBS) as a likely cause of the difference, since the 042 *aar* RBS corresponds poorly to the consensus. Hence we replaced the *aar* RBS for both the 042 and the E36 genes with the strong RBS from the pET20b expression vector to generate the pBAD30/*aar*\* 042 and pBAD30/*aar*\* E36 vectors (Table S1; Fig. S8). Biofilm formation of EAEC 042, carrying these new constructs, was then assayed, as before. Results in [Figure 7c](#) show that arabinose induction of both 042 and E36 Aar, carrying the improved RBS, completely disrupts biofilm formation. Thus, we conclude that the inability of pBAD30/*aar* 042 to interfere with biofilm formation is likely due to the poor RBS of the EAEC 042 *aar* gene, resulting in lower levels of Aar.

### Direct effects of Aar on expression from AggR-dependent promoters

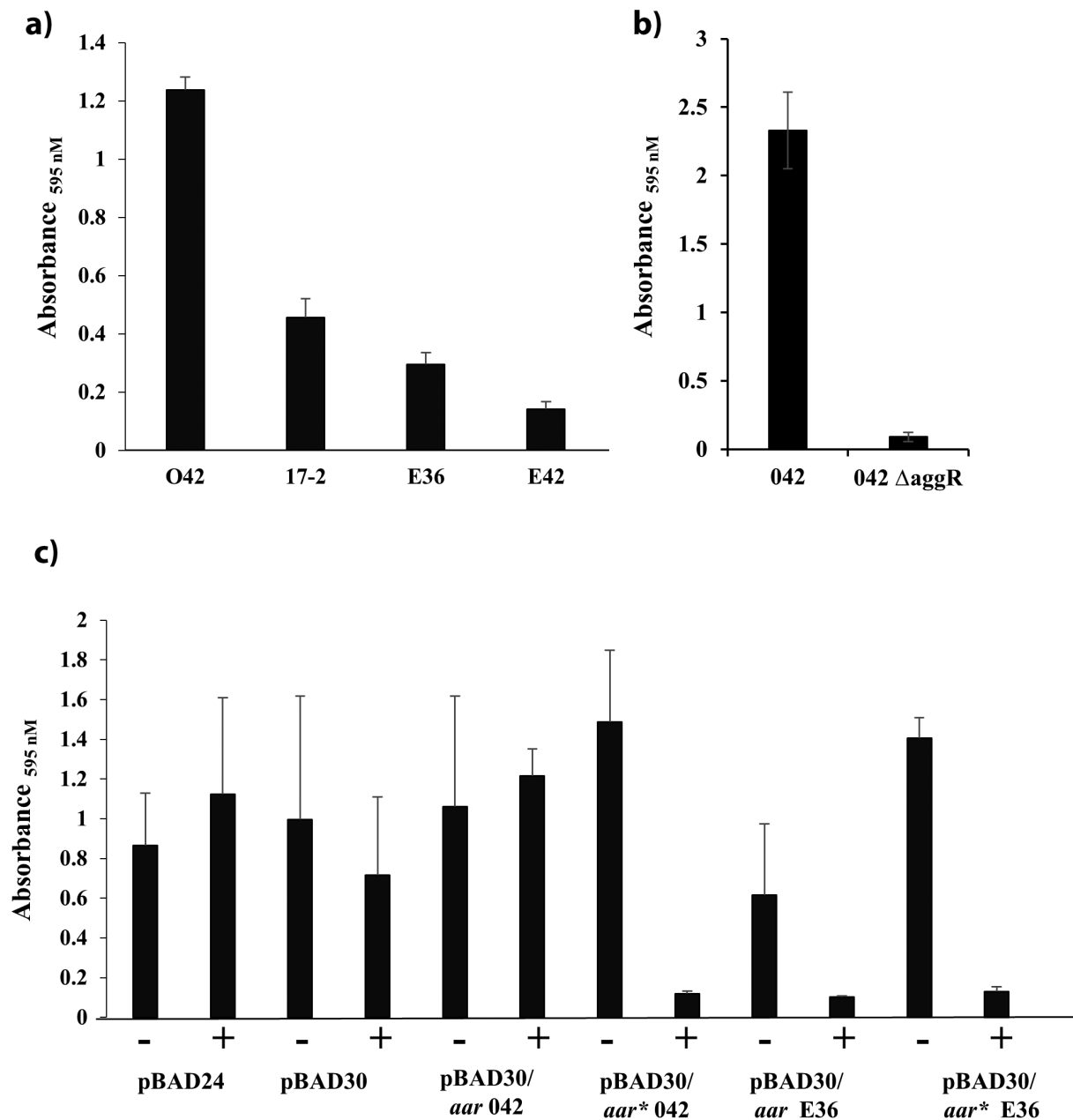
As Aar inhibits transcription activation by AggR [26,73], we used our simple two-plasmid system to examine the direct effect of Aar on AggR-dependent activation at various EAEC promoters. To achieve this, we took the pBAD/*aggR* plasmid (which carries EAEC 042 *aggR*) and cloned the DNA encoding either EAEC 042 or E36 *aar* downstream of *aggR*, to generate a synthetic operon containing both genes (*i.e.* pBAD/*aggR/aar* 042 and pBAD/*aggR/aar* E36) (Fig. S8). We also made constructs, which carry the strong RBS from pET20b (*i.e.* pBAD/*aggR/aar*\* 042 and pBAD/*aar*\* E36) (Fig. S8). These plasmids, as well as pBAD/*aggR*, were transferred into BW25113 cells, which carried a fusion of the *aggR* promoter from EAEC 042 or E36 fused to *lacZ*, as before. Cells were grown in LB medium, with 0.2% arabinose included to induce co-expression of AggR and Aar or expression of AggR alone. Surprisingly, results in [Figure 8](#) show that AggR-dependent expression from the 042 *aggR* promoter was only marginally inhibited by Aar expression, whilst that from the E36 *aggR* promoter was completely unaffected, regardless of the *aar*-containing construct used. As this may be a peculiarity of the *aggR* promoter, we also examined expression from the AggR-dependent fimbrial promoters, *afaB100* and *agg4D100* [27]. Results in [Figure 8](#) show that induction of Aar expression only decreased AggR-dependent activity from each

promoter by ~2-fold. Thus, we conclude that, although Aar expression has a considerable effect on biofilm formation, it only has a minor direct effect on expression from AggR-dependent promoters.

## Conclusions

Since their discovery, the gold standard for characterizing EAEC has been the ability of cells to adhere to HEp-2 cells in a “stacked brick” aggregative (AA) phenotype [87], with typical EAEC strains possessing *aggR* and atypical strains lacking it [20]. Recently, typical EAEC strains have been further defined molecularly as *E. coli* strains carrying *aggR* and a complete AAF gene cluster or a functional CS22 ETEC colonization factor operon [88]. Thus, it is clear that AggR-dependent activation of virulence determinant expression is the hallmark of typical EAEC strains. Our data from two new EAEC strains from Assiut, Egypt, suggest that a feed-forward activation mechanism is shared between strains from different sources. The mechanism, illustrated in ([Figure 9](#)) suggests that virulence, triggered either by a specific signal or by random fluctuations in AggR levels, follows this feed-forward activation of AggR expression. In any strain, the probability of switching into the virulent state will depend on the “basal” level of AggR, which will depend on AggR-independent activity of the *aggR* promoter, and also Aar levels, which, in turn, depend on AggR-independent *aar* promoter expression, and also the translation of Aar from the *aar* gene messenger. Our data suggest that these parameters vary from one strain to another and this may well explain the prevalence of EAEC in asymptomatic individuals: hence, the onset of virulence could well be a rare event. Note that *aar* is not present in all EAEC strains and appears to be an elaboration to the feed-forward mechanism, operating ubiquitously both pre – and post-triggering. Boisen *et al* [88], recently presented the most comprehensive survey to-date of genome sequences from nearly 100 *E. coli* strains, collected from diverse geographic settings. Our findings, reported here, are consistent with the view of Boisen *et al.* [88] of ongoing random shuffling and redistribution of different genetic determinants, which results in massive mosaic variation in EAEC strains.

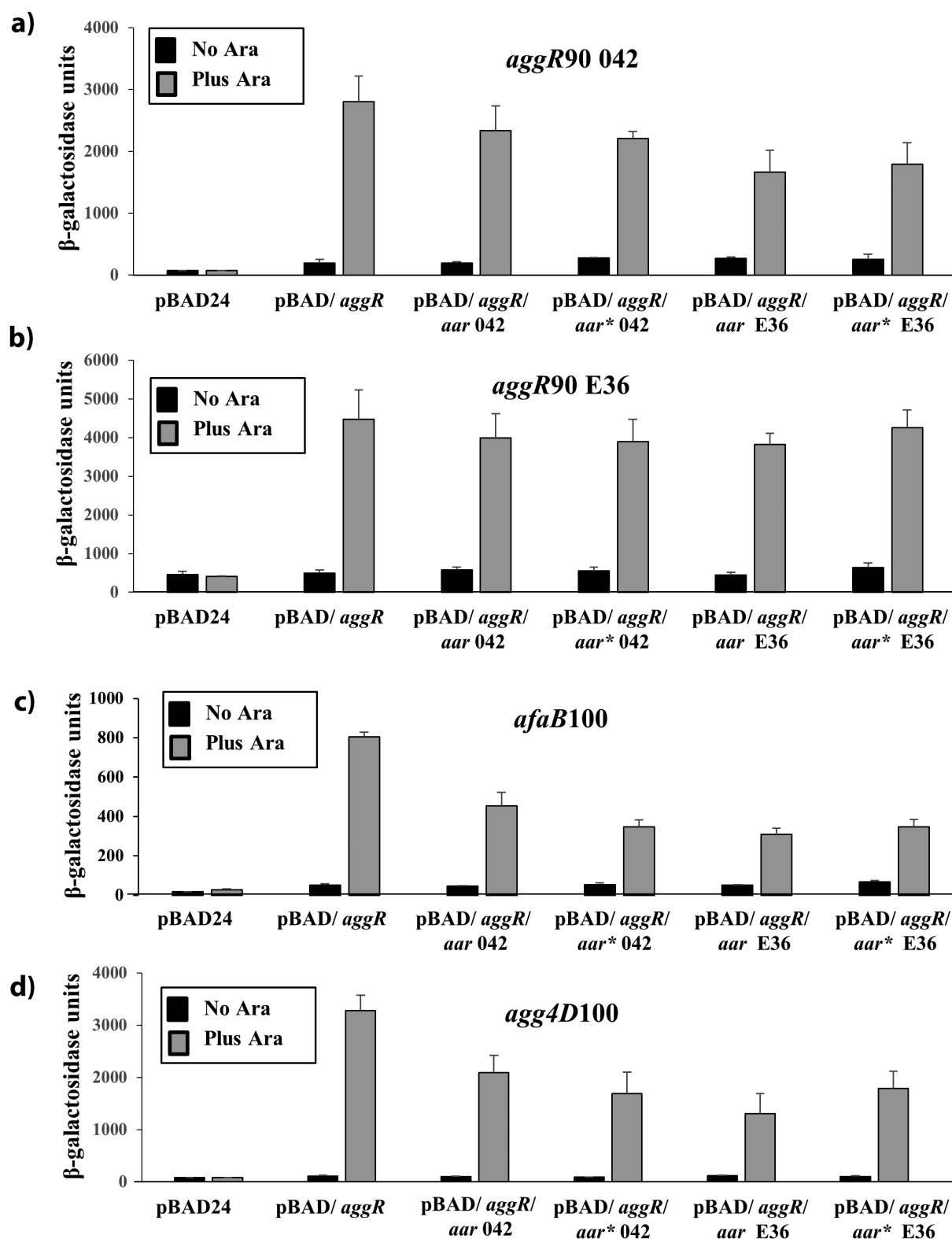
Here, we isolated *Escherichia coli* strains from infants and children, presenting with diarrhea at the outpatients clinic of Assiut University Children’s Hospital. Worryingly, all isolates were resistant to multiple classes of antibiotics and four were resistant to all antibiotics tested, highlighting the possibility that treatment with frontline antibiotics, if necessary, would



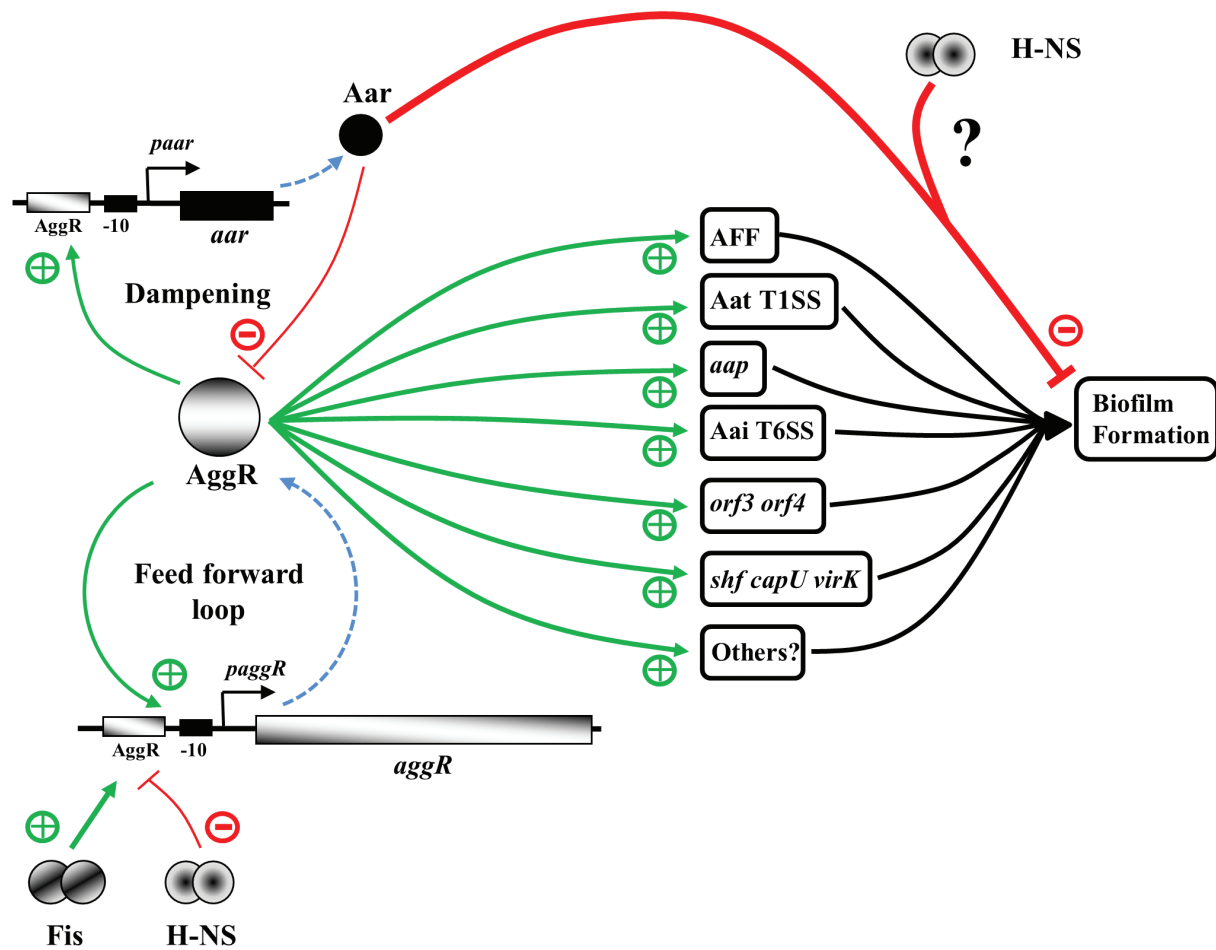
**Figure 7.** Expression of Aar inhibits biofilm formation in EAEC strain 042. A) The panel shows crystal violet biofilm formation assays, which monitor the ability of EAEC strains 042, 17-2, E36 and E42 to form biofilms on plastic microtitre plates, when grown in DMEM high glucose. B) The panel shows biofilm formation assays, monitoring the ability of EAEC 042 and EAEC 042  $\Delta$ aggR to form biofilms on plastic microtitre plates, when grown in DMEM high glucose. C) The panel shows the biofilm formed by EAEC strain 042 carrying either pBAD24, pBAD30, pBAD30/*aar* 042, pBAD30/*aar*\* 042, pBAD30/*aar* E36 or pBAD30/*aar*\* E36. Cells were grown in DMEM high glucose in presence (+) or absence (-) of 0.2% (w/v) arabinose. In all panels, the data displayed are representative experiments, with each value being an average of eight replicate samples.

likely fail to resolve many of these infections. Furthermore, all isolates were resistant to the cephalosporin antibiotics, cefaclor and ceftriaxone, suggesting that the carriage of extended spectrum  $\beta$ -lactamases (ESBL) is wide spread [89,90]. In support of this, genome analysis of E36 and E42 indicated that each strain

carried two types of ESBLs (*i.e.* *bla*<sub>CTX-M</sub> and *bla*<sub>TEM</sub>) (Table 2). We also detected other antibiotic resistance genes carried by the E36 and E42 strains (Table 2) and these likely account for their multidrug resistance phenotype. Interestingly, E36 was one of the few isolates resistant to the carbapenem antibiotics, imipenem and



**Figure 8.** The effects of Aar expression on AggR-dependent activation of target promoters. The figure shows  $\beta$ -galactosidase activities measured in *E. coli* K-12 BW25113 cells, containing *lacZ* expression vectors carrying either the A) *aggR90 042*, B) *aggR90 E36*, C) *afaB100* or D) *agg4D100* promoter fragments. In panels A) and B), promoter fragments were cloned into pRW224 and in C) and D), they were cloned into pRW50. Cells also carry either pBAD24, pBAD/*aggR*, pBAD/*aggR/aar 042*, pBAD/*aggR/aar\* 042*, pBAD/*aggR/aar E36* or pBAD/*aggR/aar\* E36*. Cells were grown in LB medium in the presence or absence of 0.2% (w/v) arabinose.  $\beta$ -galactosidase activities are expressed as nmol of ONPG hydrolyzed min<sup>-1</sup> mg<sup>-1</sup> dry cell mass. Each activity is the average of three independent determinations and standard deviations are shown for all data points.



**Figure 9.** The AggR-dependent virulence switch: a feed forward loop with dampening. In the absence of triggering, AggR-independent expression from the *aggR* and *aar* promoters maintains a balanced level of AggR and Aar, keeping the virulence switch in an off state. AggR expression is controlled positively (+) by Fis and repressed (-) by H-NS [25]. Triggering of the AggR-dependent virulence switch, either by a specific signal or a stochastic event, leads to the increased expression of AggR, amplifying the response (*i.e.* a feed forward loop), and the production of specific virulence determinants involved in host attachment, biofilm formation and protein secretion. Aar expression is also increased, inhibiting AggR-dependent transcription activation, dampening the AggR feed forward loop, and interfering with biofilm formation, possibly in concert with H-NS at specific promoters. The virulence switch will be reset by the removal of the triggering signal and Aar-mediated repression of AggR. Green arrows denote activation (+), red lines repression (-), and dotted blue arrows indicate the processes of transcription and translation for *aggR* and *aar*.

meropenem, (Table 1 and S1), and yet we did not detect any obvious carbapenem resistance genes in its genome. As carbapenem resistance can be due to a variety of mechanisms [90,91], it is possible that we failed to detect such genes due to the draft nature of our genome assemblies. Further studies will be required to determine the mechanism of carbapenem resistance in this strain.

In many studies, EAEC strains have been isolated from asymptomatic individuals [5,16]. Like E36, E42 and many pathogenic EAEC strains, 17-2 possesses a battery of AggR-regulated virulence determinants and toxins, such as the EAST-1 toxin (*astA*) [30] and the Sat secreted autotransporter toxin (Table 2) [92].

EAEC 17-2 belongs to the ST10 clonal complex, which is a major sequence type complex associated with EAEC and linked to disease causation [64]. Thus, EAEC 17-2 possesses all the characteristics of a pathogenic EAEC strain and yet, in a volunteer study, 17-2 failed to elicit diarrhea in adult volunteers [29] and, in another, only one adult out of 19 produced diarrhea [65]. Like our Egyptian EAEC strains, EAEC 17-2 was isolated from an infant with diarrhea [93]. As immune protection to EAEC strains, such as 17-2, may be acquired early in life, it has been suggested that virulence should not be determined in adult volunteers [64,65]. Thus, it is clear that, for EAEC infections, many factors contribute to whether disease occurs,

such as bacterial genotype and, importantly, the immune status, nutritional status and physical status of the infected individual [29]. It is plausible that disease is only observed in naive individuals, or when a new or particularly potent combination of virulence genes occurs, as was observed for the Shiga-toxin-producing EAEC O104:H4 outbreak strain in Germany (Table 2) [12–16,58]. Given the importance of the precise levels of AggR and Aar in cells [28], we suggest that stochastic cell-to-cell variation influences infectivity as much as genotype.

## Author Contributions

DFB, SJWB, EA, EH, NHA and ED conceived and designed the research project. RA, MY, SJE, GSC, FS, REG, NHA and DFB performed the experiments. DFB, SJWB and RA analysed the data and wrote the manuscript with input from all authors.

## Acknowledgments

This work was generously supported by a studentship from the Egyptian Ministry of Higher Education (Cultural Affairs and Missions Sector) and the Grant Office from the Medical School, Assiut University to RA, a Darwin Trust of Edinburgh PhD studentship to MY and BBSRC research grant BB/R017689/1 to DFB and SJWB. We thank MicrobesNG and Emily Jane Richardson for sequencing, genome annotation and helpful suggestions for the analysis of this data.

## Disclosure statement

The authors declare no conflict of interest.

## Funding

This work was supported by the Biotechnology and Biological Sciences Research Council [BB/R017689/1]; Egyptian Ministry of Higher Education (Cultural Affairs and Missions Sector) and the Grant Office from the Medical School, Assiut University [No Number]; Darwin Trust of Edinburgh PhD studentship [No Number].

## ORCID

Douglas F. Browning  <http://orcid.org/0000-0003-4672-3514>

## References

- [1] Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. *Nat Rev Microbiol*. 2004;2(2):123–140.
- [2] Nataro JP, Mai V, Johnson J, et al. Diarrheagenic *Escherichia coli* infection in Baltimore, Maryland, and New Haven, Connecticut. *Clin Infect Dis*. 2006;43:402–407.
- [3] Wilson A, Evans J, Chart H, et al. Characterisation of strains of enteroaggregative *Escherichia coli* isolated during the infectious intestinal disease study in England. *Eur J Epidemiol*. 2001;17(12):1125–1130.
- [4] Okeke IN, Lamikanra A, Czeizulin J, et al. Heterogeneous virulence of Enteroaggregative *Escherichia coli* strains isolated from children in Southwest Nigeria. *J Infect Dis*. 2000;181(1):252–260.
- [5] Franca FL, Wells TJ, Browning DF, et al. Genotypic and phenotypic characterisation of enteroaggregative *Escherichia coli* from children in Rio de Janeiro, Brazil. *PLoS One*. 2013;8:e69971.
- [6] Adachi JA, Jiang ZD, Mathewson JJ, et al. Enteroaggregative *Escherichia coli* as a major Etiologic agent in Traveler's Diarrhea in 3 regions of the world. *Clin Infect Dis*. 2001;32(12):1706–1709. .
- [7] Durrer P, Zbinden R, Fleisch F, et al. Intestinal infection due to Enteroaggregative *Escherichia coli* among human immunodeficiency virus-infected persons. *J Infect Dis*. 2000;182(5):1540–1544. .
- [8] Olesen B, Scheutz F, Andersen RL, et al. Enteroaggregative *Escherichia coli* O78:H10, the cause of an outbreak of urinary tract infection. *J Clin Microbiol*. 2012;50(11):3703–3711. .
- [9] Herzog K, Engeler Dusel J, Hugentobler M, et al. Diarrheagenic enteroaggregative *Escherichia coli* causing urinary tract infection and bacteremia leading to sepsis. *Infection*. 2014;42(2):441–444. .
- [10] Harrington SM, Dudley EG, Nataro JP. Pathogenesis of enteroaggregative *Escherichia coli* infection. *FEMS Microbiol Lett*. 2006;254(1):12–18.
- [11] Steiner TS, Lima AA, Nataro JP, et al. Enteroaggregative *Escherichia coli* produce intestinal inflammation and growth impairment and cause interleukin-8 release from intestinal epithelial cells. *J Infect Dis*. 1998;177(1):88–96.
- [12] Itoh Y, Nagano I, Kunishima M, et al. Laboratory investigation of enteroaggregative *Escherichia coli* O untypeable:H10 associated with a massive outbreak of gastrointestinal illness. *J Clin Microbiol*. 1997;35(10):2546–2550.
- [13] Harada T, Hiroi M, Kawamori F, et al. A food poisoning diarrhea outbreak caused by enteroaggregative *Escherichia coli* serogroup O126: h27in Shizuoka, Japan. *Jpn J Infect Dis*. 2007;60:154–155.
- [14] Frank C, Werber D, Cramer JP, et al. Epidemic profile of Shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany. *N Engl J Med*. 2011;365(19):1771–1780. .
- [15] Boisen N, Melton-Celsa AR, Scheutz F, et al. Shiga toxin 2a and Enteroaggregative *Escherichia coli* – a deadly combination. *Gut Microbes*. 2015;6(4):272–278.
- [16] Estrada-Garcia T, Navarro-Garcia F. Enteroaggregative *Escherichia coli* pathotype: a genetically heterogeneous emerging foodborne enteropathogen. *FEMS Immunol Med Microbiol*. 2012;66(3):281–298.
- [17] Henderson IR, Hicks S, Navarro-Garcia F, et al. Involvement of the Enteroaggregative *Escherichia coli*



- plasmid-encoded toxin in causing human intestinal damage. *Infect Immun.* 1999;67(10):5338–5344.
- [18] Savarino SJ, Fasano A, Robertson DC, et al. Enterotoxigenic *Escherichia coli* elaborate a heat-stable enterotoxin demonstrable in an *in vitro* rabbit intestinal model. *J Clin Invest.* 1991;87(4):1450–1455.
- [19] Nataro JP, Yikang D, Yingkang D, et al. AggR, a transcriptional activator of aggregative adherence fimbria I expression in enterotoxigenic *Escherichia coli*. *J Bacteriol.* 1994;176(15):4691–4699.
- [20] Sarantuya J, Nishi J, Wakimoto N, et al. Typical enterotoxigenic *Escherichia coli* is the most prevalent pathotype among *E. coli* strains causing diarrhea in Mongolian children. *J Clin Microbiol.* 2004;42(1):133–139.
- [21] Elias WP Jr., Czczulin JR, Henderson IR, et al. Organization of biogenesis genes for aggregative adherence fimbria II defines a virulence gene cluster in enterotoxigenic *Escherichia coli*. *J Bacteriol.* 1999;181(6):1779–1785.
- [22] Sheikh J, Czczulin JR, Harrington S, et al. A novel dispersin protein in enterotoxigenic *Escherichia coli*. *J Clin Invest.* 2002;110(9):1329–1337.
- [23] Nishi J, Sheikh J, Mizuguchi K, et al. The export of coat protein from Enterotoxigenic *Escherichia coli* by a specific ATP-binding cassette transporter system. *J Biol Chem.* 2003;278(46):45680–45689.
- [24] Morin N, Santiago AE, Ernst RK, et al. Characterization of the AggR regulon in Enterotoxigenic *Escherichia coli*. *Infect Immun.* 2013;81(1):122–132.
- [25] Morin N, Tirling C, Ivison SM, et al. Autoactivation of the AggR regulator of enterotoxigenic *Escherichia coli* in vitro and in vivo. *FEMS Immunol Med Microbiol.* 2010;58(3):344–355.
- [26] Santiago AE, Ruiz-Perez F, Jo NY, et al. A large family of antivirulence regulators modulates the effects of transcriptional activators in gram-negative pathogenic bacteria. *PLoS Pathog.* 2014;10:e1004153.
- [27] Yasir M, Icke C, Abdelwahab R, et al. Organization and architecture of AggR-dependent promoters from enterotoxigenic *Escherichia coli*. *Mol Microbiol.* 2019;111(2):534–551.
- [28] Mickey AS, Nataro JP. Dual function of Aar, a member of the new AraC negative regulator family, in *Escherichia coli* gene expression. *Infect Immun.* 2020;88(6). DOI:10.1128/IAI.00100-20
- [29] Nataro JP, Deng Y, Cookson S, et al. Heterogeneity of enterotoxigenic *Escherichia coli* virulence demonstrated in volunteers. *J Infect Dis.* 1995;171(2):465–468.
- [30] Veilleux S, Holt N, Schultz BD, et al. *Escherichia coli* EAST1 toxin toxicity of variants 17-2 and O 42. *Comp Immunol Microbiol Infect Dis.* 2008;31:567–578.
- [31] Zhang R, Gu DX, Huang YL, et al. Comparative genetic characterization of Enterotoxigenic *Escherichia coli* strains recovered from clinical and non-clinical settings. *Sci Rep.* 2016;6(1):24321.
- [32] Guiral E, Mendez-Arancibia E, Soto SM, et al. CTX-M-15-producing Enterotoxigenic *Escherichia coli* as cause of travelers' Diarrhea. *Emerg Infect Dis.* 2011;17(10):1950–1953.
- [33] Chattaway MA, Jenkins C, Ciesielczuk H, et al. Evidence of evolving Extraintestinal Enterotoxigenic *Escherichia coli* ST38 Clone. *Emerg Infect Dis.* 2014;20(11):1935–1937.
- [34] Ali MM, Ahmed SF, Klena JD, et al. Enterotoxigenic *Escherichia coli* in diarrheic children in Egypt: molecular characterization and antimicrobial susceptibility. *J Infect Dev Ctries.* 2014;8(05):589–596.
- [35] Sanders JW, Putnam SD, Gould P, et al. Diarrheal illness among deployed U.S. military personnel during Operation Bright Star 2001–Egypt. *Diagn Microbiol Infect Dis.* 2005;52(2):85–90.
- [36] Riddle MS, Halvorson HA, Shiao D, et al. Acute gastrointestinal infection, respiratory illness, and non-combat injury among US military personnel during Operation Bright Star 2005, in Northern Egypt. *J Travel Med.* 2007;14(6):392–401.
- [37] Meheissen M, Header D, Abdelaty K. Phylogenetic and pathotype analysis of *Escherichia coli* stool isolates from Egyptian patients with inflammatory bowel disease. *Germs.* 2019;9(4):172–181.
- [38] Ahmed SF, Shaheen HI, Abdel-Messih IA, et al. The epidemiological and clinical characteristics of diarrhea associated with enteropathogenic, enterotoxigenic and diffuse-adherent *Escherichia coli* in Egyptian children. *J Trop Pediatr.* 2014;60(5):397–400.
- [39] Sambrook J, Russell DW. *Molecular cloning: a laboratory manual.* 3rd ed. Cold Spring Harbor, N. Y: Cold Spring Harbor Laboratory Press; 2001.
- [40] Lodge J, Fear J, Busby S, et al. Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS Microbiol Lett.* 1992;74(2–3):271–276.
- [41] Islam MS, Bingle LE, Pallen MJ, et al. Organization of the LEE1 operon regulatory region of enterohaemorrhagic *Escherichia coli* O157: H7 O157:H7 and activation by GrlA. *Mol Microbiol.* 2011;79(2):468–483.
- [42] Bauer AW, Kirby WM, Sherris JC, et al. Antibiotic susceptibility testing by a standardized single disk method. *Am J Clin Pathol.* 1966;45:493–496.
- [43] Clinical\_and\_Laboratory\_Standards\_Institute\_(CLSI). Performance standards for antimicrobial susceptibility testing; Twenty-Fourth informational supplement, CLSI Document M100-S24, Wayne. 2014;34(1)
- [44] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–2120.
- [45] Wick RR, Judd LM, Gorrie CL, et al. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol.* 2017;13(6):e1005595.
- [46] Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014;30(14):2068–2069.
- [47] Rutherford K, Parkhill J, Crook J, et al. Artemis: sequence visualization and annotation. *Bioinformatics.* 2000;16(10):944–945.
- [48] Grant JR, Stothard P. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Res.* 2008;36(Web Server):W181–4.
- [49] Carver TJ, Rutherford KM, Berriman M, et al. ACT: the Artemis Comparison Tool. *Bioinformatics.* 2005;21(16):3422–3423.

- [50] Carver T, Thomson N, Bleasby A, et al. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics*. 2009;25(1):119–120.
- [51] Larsen MV, Cosentino S, Rasmussen S, et al. Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol*. 2012;50(4):1355–1361. .
- [52] Joensen KG, Tetzschner AM, Iguchi A, et al. Rapid and easy in Silico serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. *J Clin Microbiol*. 2015;53(8):2410–2426.
- [53] Carattoli A, Zankari E, Garcia-Fernandez A, et al. *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother*. 2014;58:3895–3903.
- [54] Zankari E, Hasman H, Cosentino S, et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother*. 2012;67:2640–2644.
- [55] Joensen KG, Scheutz F, Lund O, et al. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *J Clin Microbiol*. 2014;52(5):1501–1510. .
- [56] Zhang Z, Schwartz S, Wagner L, et al. A greedy algorithm for aligning DNA sequences. *J Comput Biol*. 2000;7(1–2):203–214.
- [57] Dunne KA, Chaudhuri RR, Rossiter AE, et al. Sequencing a piece of history: complete genome sequence of the original *Escherichia coli* strain. *Microb Genom*. 2017;3:mgen000106.
- [58] Rasko DA, Webster DR, Sahl JW, et al. Origins of the *E. coli* strain causing an outbreak of Hemolytic–Uremic Syndrome in Germany. *N Engl J Med*. 2011;365(8):709–717. .
- [59] Page AJ, Cummins CA, Hunt M, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. 2015;31(22):3691–3693. .
- [60] Stamatakis A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30(9):1312–1313.
- [61] Sarkar G, Sommer SS. The “megaprimer” method of site-directed mutagenesis. *Biotechniques*. 1990;8:404–407.
- [62] Guzman LM, Belin D, Carson MJ, et al. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol*. 1995;177(14):4121–4130.
- [63] Browning DF, Lee DJ, Wolfe AJ, et al. The *Escherichia coli* K-12 NarL and NarP proteins insulate the *nrf* promoter from the effects of integration host factor. *J Bacteriol*. 2006;188(21):7449–7456.
- [64] Okeke IN, Wallace-Gadsden F, Simons HR, et al. Multi-locus sequence typing of enteroaggregative *Escherichia coli* isolates from Nigerian children uncovers multiple lineages. *PLoS One*. 2010;5(11):e14093. .
- [65] Nataro JP, Deng Y, Maneval DR, et al. Aggregative adherence fimbriae I of enteroaggregative *Escherichia coli* mediate adherence to HEp-2 cells and hemagglutination of human erythrocytes. *Infect Immun*. 1992;60(6):2297–2304.
- [66] Villa L, Garcia-Fernandez A, Fortini D, et al. Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. *J Antimicrob Chemother*. 2010;65(12):2518–2529.
- [67] Chaudhuri RR, Sebaihia M, Hobman JL, et al. Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative *Escherichia coli* strain 042. *PLoS One*. 2010;5(1):e8801. .
- [68] Ho PL, Lo WU, Lai EL, et al. Clonal diversity of CTX-M-producing, multidrug-resistant *Escherichia coli* from rodents. *J Med Microbiol*. 2015;64(2):185–190. .
- [69] Scholz P, Haring V, Wittmann-Liebold B, et al. Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010. *Gene*. 1989;75:271–288.
- [70] Timofte D, Maciucă IE, Evans NJ, et al. Detection and molecular characterization of *Escherichia coli* CTX-M-15 and *Klebsiella pneumoniae* SHV-12 beta-lactamases from bovine mastitis isolates in the United Kingdom. *Antimicrob Agents Chemother*. 2014;58(2):789–794. .
- [71] Zhang L, Lu X, Zong Z. The emergence of blaCTX-M-15-carrying *Escherichia coli* of ST131 and new sequence types in Western China. *Ann Clin Microbiol Antimicrob*. 2013;12(1):35.
- [72] Smet A, Van Nieuwerburgh F, Vandekerckhove TT, et al. Complete nucleotide sequence of CTX-M-15-plasmids from clinical *Escherichia coli* isolates: insertional events of transposons and insertion sequences. *PLoS One*. 2010;5(6):e11202. .
- [73] Santiago AE, Yan MB, Tran M, et al. A large family of anti-activators accompanying XylS/AraC family regulatory proteins. *Mol Microbiol*. 2016;101:314–332.
- [74] Jonsson R, Struve C, Boisen N, et al. Novel aggregative adherence fimbria variant of enteroaggregative *Escherichia coli*. *Infect Immun*. 2015;83(4):1396–1405. .
- [75] Boisen N, Struve C, Scheutz F, et al. New adhesin of enteroaggregative *Escherichia coli* related to the Afa/Dr/AAF family. *Infect Immun*. 2008;76(7):3281–3292.
- [76] Fujiyama R, Nishi J, Imuta N, et al. The *shf* gene of a *Shigella flexneri* homologue on the virulent plasmid pAA2 of enteroaggregative *Escherichia coli* 042 is required for firm biofilm formation. *Curr Microbiol*. 2008;56:474–480.
- [77] Dudley EG, Thomson NR, Parkhill J, et al. Proteomic and microarray characterization of the AggR regulon identifies a *pheU* pathogenicity island in enteroaggregative *Escherichia coli*. *Mol Microbiol*. 2006;61:1267–1282.
- [78] Sheikh J, Dudley EG, Sui B, et al. EilA, a HilA-like regulator in enteroaggregative *Escherichia coli*. *Mol Microbiol*. 2006;61(2):338–350.
- [79] Lund P, Tramonti A, De Biase D. Coping with low pH: molecular strategies in neutralophilic bacteria. *FEMS Microbiol Rev*. 2014;38:1091–1125.
- [80] Johnson TJ, Wannemuehler YM, Nolan LK. Evolution of the *iss* gene in *Escherichia coli*. *Appl Environ Microbiol*. 2008;74:2360–2369.
- [81] Clermont O, Bonacorsi S, Bingen E. Rapid and simple determination of the *Escherichia coli* Phylogenetic

- Group. *Appl Environ Microbiol.* **2000**;66(10):4555–4558.
- [82] Toma C, Higa N, Iyoda S, et al. The long polar fimbriae genes identified in Shiga toxin-producing *Escherichia coli* are present in other diarrheagenic *E. coli* and in the standard *E. coli* collection of reference (ECOR) strains. *Res Microbiol.* **2006**;157(2):153–161.
- [83] Benjelloun-Touimi Z, Sansonetti PJ, Parsot C. SepA, the major extracellular protein of *Shigella flexneri*: autonomous secretion and involvement in tissue invasion. *Mol Microbiol.* **1995**;17(1):123–135.
- [84] Browning DF, Godfrey RE, Richards KL, et al. Exploitation of the *Escherichia coli lac* operon promoter for controlled recombinant protein production. *Biochem Soc Trans.* **2019**;47:755–763.
- [85] Sheikh J, Hicks S, Dall’Agnol M, et al. Roles for Fis and YafK in biofilm formation by enteroaggregative *Escherichia coli*. *Mol Microbiol.* **2001**;41(5):983–997.
- [86] Boisen N, Hansen AM, Melton-Celsa AR, et al. The Presence of the pAA Plasmid in the German O104: h4Shiga O104:H4 Shiga Toxin Type 2a (Stx2a)-producing Stx2a)-Producing Enteroaggregative *Escherichia coli* strain promotes the translocation of Stx2a across an Epithelial cell Monolayer. *J Infect Dis.* **2014**;210(12):1909–1919. .
- [87] Nataro JP, Kaper JB, Robins-Browne R, et al. Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells. *Pediatr Infect Dis J.* **1987**;6(9):829–831.
- [88] Boisen N, Østerlund MT, Joensen KG, et al. Redefining enteroaggregative *Escherichia coli* (EAEC): genomic characterization of epidemiological EAEC strains. *PLoS Negl Trop Dis.* **2020**;14(9):e0008613. .
- [89] Oteo J, Perez-Vazquez M, Campos J. Extended-spectrum [beta]-lactamase  $\beta$ -lactamase producing *Escherichia coli*: changing epidemiology and clinical impact. *Curr Opin Infect Dis.* **2010**;23(4):320–326.
- [90] Tooke CL, Hinchliffe P, Bragginton EC, et al. beta-Lactamases and beta-Lactamase Inhibitors in the 21st Century. *J Mol Biol.* **2019**;431(18):3472–3500. .
- [91] Chetri S, Bhowmik D, Paul D, et al. AcrAB-TolC efflux pump system plays a role in carbapenem non-susceptibility in *Escherichia coli*. *BMC Microbiol.* **2019**;19:210.
- [92] Guyer DM, Henderson IR, Nataro JP, et al. Identification of sat, an autotransporter toxin produced by uropathogenic *Escherichia coli*. *Mol Microbiol.* **2000**;38:53–66.
- [93] Vial PA, Robins-Browne R, Lior H, et al. Characterization of enteroadherent-aggregative *Escherichia coli*, a putative agent of diarrheal disease. *J Infect Dis.* **1988**;158(1):70–79. .