ROMA TRE
UNIVERSITÀ DEGLI STUDI

**DOCTORAL SCHOOL IN BIOLOGY**

Section: Biology applied to human health

**CYCLE XXV**

# Identification and functional analysis of virulence determinants of Verocytotoxin-producing *Escherichia coli* (VTEC)

**Valeria Michelacci**
MSc. (Hons), Genomic Biotechnology

**Tutor**
Dr. Stefano Morabito

**Coordinator**
Prof. Paolo Visca

**External Referees**
Prof. Herbert Schmidt (Ext. Examiner)
Prof. Maria Teresa Muniesa Perez
Dr. Patrick Fach

Thesis submitted for the Degree of Doctor of Philosophy
December 2012

I confirm that the thesis is within the size limits as stated by the University of Roma Tre rules

# Abstract

Verocytotoxin(VT)-producing *Escherichia coli* (VTEC) are important zoonotic pathogens whose natural reservoir is the gastrointestinal tract of ruminants. The transmission of the infections mainly occurs *via* the ingestion of contaminated food of animal origin. VTEC pathogenicity relies on the production of the VTs and on the action of accessory virulence factors constituting the virulome, which has not been completely identified yet.

The main objective of this piece of research was the identification of the genomic structures forming the VTEC virulome. An additional goal was the analysis of their distribution in different VTEC sub-populations. Finally, a function for some of the factors identified and involved in the pathogenetic mechanism has been proposed.

The work presented here has been largely based on the genomic comparison of VTEC strains isolated from human and animal sources and held in the collections of the EU RL VTEC and of the collaborating institutions Statens Serum Institut (Copenhagen, DK) and University of Extremadura (Caceres, ES). This approach led to the identification and characterisation of two pathogenicity islands (PAIs) proposed to be part of the virulome of VTEC strains commonly isolated from cases of human disease. The two PAIs harbour genes encoding factors involved in the colonization (*adfO*, OI-57, Chapter 4 and *tia*, SE-PAI, Chapter 3) or specifying an allelic variant of the Subtilase cytotoxin (s*ubAB*$_2$, SE-PAI, Chapter 3).

The identification and characterization of the open reading frames carried by the two PAIs allowed making inference on the function of the encoded proteins and on their role in the pathogenetic process.

As a matter of fact, the *tia* and *shiA* genes (SE-PAI, Chapter 3) have been proposed to be part of the colonization machinery of VTEC strains lacking the ability to cause *attaching-and-effacing* (A/E) lesions on the intestinal mucosa, a typical feature of the VTEC associated with the most severe forms of the infection. The products of genes homologues to *tia* and *shiA* have been in fact described respectively to have a role in the process of cell invasion used by some enterotoxigenic *E. coli* strains or in the attenuation of host inflammatory response upon infection of another invasive enteric pathogen: *Shigella flexneri*.

The genomic approach has also been used to investigate on the VT-producing Enteroaggregative *E. coli* O104:H4 that caused an outbreak in Germany in 2011. This peculiar VTEC is characterised by a rare combination of virulence traits from Enteroaggregative *E. coli* (EAEC) and VTEC, two different *E. coli* pathotypes responsible for enteric infections in low-income areas and common in industrialized countries respectively. Before the German outbreak, VTEC O104:H4 had been isolated in a few sporadic cases of infection, sometimes with an epidemiological connection with travel in North Africa. Similarly, the strain that caused the German outbreak was introduced into the EU with fenugreek seeds produced in Egypt. These observations, together with the genomic stability observed for some of the VTEC O104 strain (Chapter 5), led to the hypothesis that such chimeric strains may have emerged in developing countries. The high loads of enteric pathogens causing diarrhoea and lack of wastewater treatment in these geographic regions, cause the diffuse faecal contamination of the environment, particularly of the aquatic ecosystem often shared with ruminants, facilitating the exchange of mobile genetic elements between human and animal *E. coli* pathotypes and leading to the formation of new combinations of virulence features such as those characterising the VT-producing EAEC O104:H4.

## Acknowledgements

This PhD thesis is the outcome of the last three years spent at the Department of Food Safety and Veterinary Public Health of the Istituto Superiore di Sanità (ISS) in Rome: a period of highly valuable professional and human experience in a widely international atmosphere and during which times of frustration have always been swept away by exciting and rewarding moments.

I am indebted to many people for making the time working on my PhD an unforgettable experience

My first debt of gratitude goes to my advisor, Dr Stefano Morabito, who patiently provided the vision, encouragement and advise necessary for me to proceed through the doctorial programme and complete my dissertation. My best wish for my future would be to be able to command an audience someday as well as he can.

I am also very grateful to Dr Alfredo Caprioli for his constant and prompt availability in discussing scientific questions and giving interesting suggestions especially in revising papers or oral presentations, no matter how busy he could be.

I owe a lot to my dear friend and supervisor Rosy, who has worked together with me during all the period and has become busier and busier while proceeding with her scientific career, but who has never regret to help me in any way she could, either discussing about my scientific doubts or helping with any kind of benchwork or finally supporting me in my personal affairs.

I had the wonderful opportunity to spend several periods as a host in outstanding foreign laboratories, where I was always so lucky to meet other great scientists taking care about me with the same kindness I was used to at home. In this respect, it has been a real honor and pleasure to me to collaborate with Prof. Pina Fratamico at the U.S. Department of Agriculture in Pennsylvania, who has acted as a chief and a mother at the same time, making me feel completely at home even at the other edge of the world; Dr Flemming Scheutz at the Statens Serum Institut in Copenhagen, who I have had the opportunity to meet, work with and discuss with in a number of occasions during these years and always demonstrated kindness and availability to me, as only great people like him can do; and Prof Patrick Fach at ANSES in Maison-Alfort, who gave me the opportunity to "play" with some of the most futuristic RealTime-PCR-machines I had ever heard about. Many thanks also go to all the other people who assisted me kindly and patiently during my stages: David Needleman at USDA, who patiently helped me during my stage at USDA; Susanne and Pia, who provided me with any kind of help during the long and cold periods spent in Copenhagen; and Sabine Dellanoy at ANSES, who dealt with the analysis of a huge amount of data together with me, helping me not to get lost in the resulting excel files.

Many thanks also to Prof. Paolo Visca for acting as a supervisor during my PhD and to Prof Schmidt, Dr Muniesa and Prof Fach for accepting to review my thesis.

Special thanks also go to my friends and colleagues at ISS: Maria Luisa, Clarissa, Fabio, Paola and especially Laura and Antonella, who have been helping me in the lab work and have supported me especially in the last year of my PhD programme, when I traveled so much that I almost forgot where I was coming from. They, and Rosy as well, were always ready to remind me that I had nice people in the lab in Rome willing to help me in any way, even when they were facing hard periods on their own. Many other people passed by the lab for shorter periods, like Lejla and Remi, who worked a lot together with me and separately helped for the publication of two of the papers part of this thesis, besides becoming

nice friends to go out with and spend a nice time, no matter where we meet: it has already happened in Rome, Amsterdam and Copenhagen!
Thanks to my dearest friends Ilaria, Elena and Maria Letizia, who have always supported me and found a way to make me laugh even if I have often been too busy to meet them frequently.
Special thanks go to my family for their support and understanding, even in my hardest periods.
Finally, Luca wouldn't need me to write here how thankful I am to him: he has taken an active part in the most difficult as well in the nicest moments of the last years and has always been the closest person to me, no matter how far we were. He would deserve a *honoris causa* degree for the knowledge he has patiently acquired on *E. coli* while listening to me.
I strongly believe that these three years have opened my mind in many ways; now I can't wait to see what other interesting opportunities are coming up for me in the next future.

# Table of contents

# 1.Introduction

# 1.INTRODUCTION

### 1.1 *Escherichia coli*

*Escherichia coli* is a Gram-negative bacterium belonging to the family *Enterobacteriaceae* and represents an important component of the human and animal intestinal microflora.

This bacterial species colonizes the gastrointestinal tract during the first phases of the life establishing mutual beneficial relationships with the host and playing an important role in maintaining the equilibrium between the numerous bacterial species colonising the gut, in regulating the turnover of the intestinal epithelium and in promoting the development of local immunity system against pathogens. At the same time, *E. coli* is one of the most diffuse bacterial species in the environment, being present in almost all the niches including water and soil.

Some strains evolved the capability to harm and cause disease. The pathogenic variants of *E. coli* can cause a wide spectrum of diseases and are generally divided into two main large categories, or pathogroups, based on the district of the host they colonize: Extraintestinal Pathogenic *E. coli* (ExPEC) and Diarrhoeagenic *E. coli* (DEC).

ExPEC produce virulence features conferring the ability to colonize host districts other than the gastrointestinal tract. The pathogroup is divided into two main distinct populations: Uropathogenic *E. coli* (UPEC) and neonatal meningitis *E. coli* (NMEC) (Smith JL *et al*., 2007) causing a variety of human disease including urinary tract infections, neonatal meningitis, sepsis, pneumonia and surgical site infections.

Diarrhoeagenic *E. coli* (DEC) comprise different groups, or pathotypes, of *E. coli*, which infect the gut of the host inducing a wide spectrum of symptoms spanning from uncomplicated diarrhoea to haemorrhagic colitis and systemic complications like haemolytic uremic syndrome. DEC infections represent a major problem in low-income countries, where mortality from diarrhoea is one of the main causes of death every year (Kosek M. *et al*., 2003; Bern C. *et al*., 2004).

### 1.2 Genomic plasticity of *E. coli*

The exceptional capability of *E. coli* strains to colonize a broad range of hosts and environments is linked to their ability of establishing successful relationships with the other microorganisms and it is largely due to their extraordinary genomic plasticity. This is the capability to exchange genetic material with other bacteria also belonging to other species, through events of gain and loss of DNA traits.

The *E. coli* genome is shaped by a multitude of evolutionary forces derived from its habitat, in which either biotic (e. g competitors, host defence mechanisms) or abiotic (e. g. pH, temperature, UV, mineral depletion) factors determine the selection of the individuals most adapted to survive in any given niche (Van Elsas J.D. *et al*., 2011). As a result of this selective pressure, the *E. coli* genomes vary in size from 4.6 to 5.6 Mb (Bergthorsson U. & Ochman H., 1995). These differences are due to the presence of different amounts of strain-specific genetic information, which may represent up to 30% of the complete genome content (Rasko D.A. *et al*., 2008; Touchon M. *et al*. 2009, Dobrindt U. *et al.*, 2010).

Comparative genome analysis revealed that *E. coli* genomes have a mosaic-like structure consisting of a conserved part, the ‘‘core genome’’, and interspersed regions of variable DNA. The core genome is common to all the *E. coli* strains and comprises about 2,000 genes, governing the basic metabolic functions. It represents about 11% of the totality of the so-called *E. coli* “pangenome”, which includes all the *E. coli* genetic features discovered so far (Fig. 1). A large portion (62%) of the pangenome is composed of so-called ‘persistent' genes that, in different combination with the core genome, identify the different *E. coli* pathogroups. Finally, the remaining 26% of the pangenome is composed by ‘volatile' genes conferring pathotype/strain specificity (Touchon M. *et al*., 2009).

Remodelling of the pangenome resulted in the evolution of specific *E. coli* types (Van Elsas J.D. *et al.*, 2011) including both the commensal and pathogenic variants (pathotypes) (Ochman H. *et al*. 2000).

The process of exchange of DNA traits between bacteria is defined as horizontal gene transfer (HGT) and is characterized by the integration and excision of mobile genetic elements (MGE). These elements consist in DNA regions encoding all the components required for their mobilisation and able to allocate part of the genetic material deriving from the host genome, variable in size from a few to several thousand base pairs.

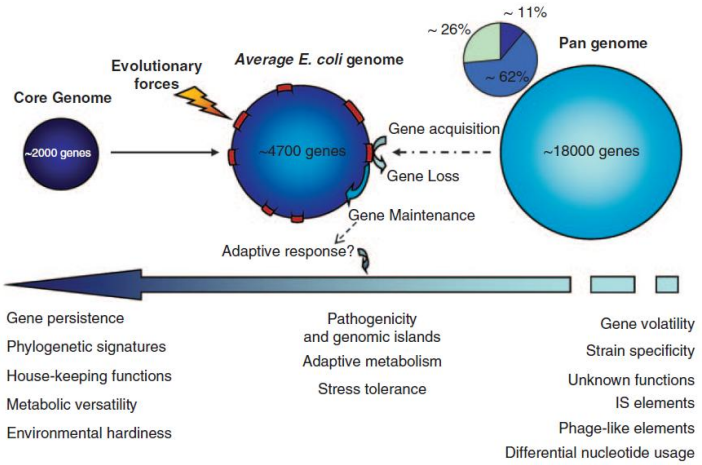



*Figure 1. The E. coli pangenome. (Van Elsas J.D. et al., 2011)*

Through HGT, bacteria can easily acquire MGEs, vehiculating genes encoding advantageous functions or loose genetic determinants encoding unwanted or obsolete characters.

The horizontal transfer of MGE represents the key for the success of *E. coli* as ubiquitous bacterial species. HGT increases the genetic variability favouring the selection of the individuals most adapted to win the “struggle for life”.

### 1.3 Mobile genetic elements

Mobile genetic elements include bacteriophages, plasmids, genomic islands, integrons, transposons and insertion sequence elements (Dobrindt U. *et al*., 2005). The first three categories play the most important role in the evolution of pathogenic *E. coli* (Kelly B.G. *et al*., 2009) and will be discussed in detail.

1.3.1 *Bacteriophages*

Bacteriophages had a profound impact on the evolution of bacterial pathogens (Campbell A., 1996). As a matter of fact, the spread of virulence-associated genes by lysogenic phages is a common phenomenon in many bacterial species (Boyd E.F. & Brussow H., 2002). Moreover, temperate and cryptic phages are considered possible ‘‘ancestors’’ of pathogenicity islands (Dobrindt U. *et al*., 2004).

The role of phages in modifying bacterial virulence was first demonstrated by Freeman in 1951, when he observed that the infection of non-toxigenic strains of *Corynebacterium diphtheria* with a bacteriophage resulted in the production of the diphtheria toxin (Freeman V.J., 1951; Tinsley C.R. *et al*., 2006). Since then, many toxins and other virulence-associated factors have been demonstrated to be encoded by genes carried by cryptic or intact phages in Gram-positive and negative bacteria (Casjens S., 2003).

The analyses of the *E. coli* genomes published so far, indicated that the sequenced strains differ considerably in the number of prophage-related sequences (Blattner F.R. *et al*., 1997; Hayashi T. *et al*., 2001; Perna N.T. *et al*., 2001; Welch R.A. *et al*., 2002) and that genome size differences correlate with the number of these MGEs, thus confirming the importance of the role of phages in structuring *E. coli* genome.

*1.3.2 Plasmids*

Plasmids are circular, self-replicating DNA molecules existing in bacterial cells as extra chromosomal replicons and are very common in all bacterial species. They represent a distinct genetic resource as they may confer many advantageous features such as resistance to antibiotics, expression of bacteriocins and siderophores, (Gomez-Lus R., 1998). Plasmids vehiculate genes playing important roles in the virulence mechanisms of pathogenic bacteria including DEC and ExPEC.

Plasmids can be easily exchanged and quickly disseminated in the bacterial populations representing one of the most characterised vehicles for the horizontal gene transfer between bacteria.

*1.3.3 Genomic Islands*

The concept of genomic islands (GI) was conceived following the discovery of particular regions in the bacterial chromosomes that contributed to pathogenicity and were termed pathogenicity islands (PAI) (Hacker J. & Kaper J.B., 2000). GI are characterized by a significant difference in G + C content, compared to the average base composition of the bacterial chromosome, an alternative codon usage, presence of mobility genes, and, in PAIs, of virulence genes. The majority of the GIs carry functions that are useful for the survival of the organism, by providing a selective advantage over the populations that do not harbour the islands (Hacker J. & Carniel E., 2001). Many types of GIs have been described, including those carrying genes governing metabolic properties enhancing the individual fitness, conferring resistance to antimicrobials or involved in the pathogenetic mechanisms (Dobrindt U. *et al*., 2004). The type of island that an organism carries depends on the genetic background of the host and on its ecological niche.

Pathogenicity islands carry one or more virulence genes, are present in the genomes of pathogenic bacteria but absent from the non-pathogenic variant of the same species and often exist in the size range of hundreds bp to more than 200 kb (Schmidt H. & Hensel M., 2004). Not much is known about the

origin of these GIs, but it is thought that they originated from co-integrated plasmids and/or phages that have lost the ability to self-transfer.

GIs are typical features of *E. coli* and played an important role for evolution of the different variants or pathotypes. Many virulence-associated genes of pathogenic *E. coli* are located on PAIs, including those encoding adhesins, toxins, iron uptake systems, secretion systems and molecules involved in inhibiting the host response.

In pathogenic *E. coli* strains up to 13 PAI-like genetic entities have been identified (Hayashi T. *et al*., 2001; Perna N.T. *et al*., 2001; Welch R.A. *et al*., 2002).

Many PAIs have a mosaic-like modular structure. They can have different structural organizations or chromosomal localizations in different strains, still vehiculating the same virulence determinants (Guyer D.M. *et al*., 1998; Johnson J.R. *et al*., 2001; Redford P. & Welch R.A., 2002). This is probably due to the presence of multiple copies of accessory DNA elements facilitating the homologous recombination within the same or between different islands leading to rearrangements, deletions and acquisition of ‘‘foreign’’ DNA.

### 1.4 Pathogroups and pathotypes of *E. coli*: Emergence and characteristics

Horizontal gene transfer has driven *E. coli* evolution resulting in the definition of specific *E. coli* lineages, comprising commensal and pathogenic variants (Ochman H. *et al*. 2000).

All *E. coli* strains were initially thought to be non-pathogenic commensal organisms. In the 1940s *E. coli* strains were first identified in association with severe outbreaks of infantile diarrhoea (Bray J., 1945). Since then, *E. coli* has been associated with a range of clinical conditions and several pathogenic groups have been defined (Kaper J.B. *et al*., 2004; Nataro J.P. & Kaper J.B., 1998).

#### *1.4.1 Extraintestinal pathogenic E. coli (ExPEC)*

*E. coli* strains isolated from infections outside of the intestinal tract, e.g., uropathogenic *E. coli* (UPEC), neonatal meningitis-associated *E. coli* (NMEC) and sepsis-causing *E. coli* (SEPEC) have been grouped as Extraintestinal pathogenic *E. coli* (ExPEC) (Smith J.L. *et al*., 2007). ExPEC are part of the intestinal microflora in a fraction of the healthy population and normally asymptomatically colonize the gut. Once they get access to niches outside of the gut, they are able to efficiently colonize these niches and cause disease, i.e. urinary tract infection (UTI), septicemia or meningitis in newborns in humans and in many animal species (Köhler C.D. & Dobrindt U., 2011). Several important virulence factors of ExPEC and their role during pathogenesis have been described (Kaper J.B. *et al*., 2004; Smith J.L. *et al*., 2007), however the complete gene asset conferring virulence has not been identified yet. As a matter of fact, ExPEC strains can use multiple virulence determinants in a mix-and-match fashion.

These virulence associated factors can be grouped by functional category, as they comprise, for example, adhesins, siderophore systems, toxins, surface polysaccharides, features involved in the invasion process and serum resistance-associated traits (Johnson J.R. & Russo T.A., 2005).

The comparison between the genomes of CFT073 UPEC strain and non-pathogenic *E. coli* strain Nissle 1917 revealed that, although the overall genetic structure of many genomic islands is very similar, important differences exist which are responsible for the non-pathogenic nature of strain Nissle 1917

(Grozdanov L. *et al*., 2004). The *pheV*-associated PAI of strain CFT073 carries the complete *hly* and *pap* gene clusters respectively encoding the alpha-hemolysin and P fimbriae virulence factors. Interestingly, only a fragmented *pap* operon is present in a similar DNA region on the chromosome of Nissle 1917 associated with a transposon-like element, suggesting that the original content of the PAI may have been lost due to the insertion of mobile genetic elements and recombination events.

Five cryptic prophages have been detected in the genome of CFT073 strain (Welch R.A. *et al*., 2002), which exhibit some structural features similar to some of the phage-like regions in the *E. coli* O157:H7 genome, besides containing different putative virulence-associated genes (Dobrindt U., 2005)

Other virulence-associated features, like the cytotoxic necrotizing factor 2 (Caprioli A. *et al*., 1987) and the characteristic F17 fimbria, are encoded by genes harboured on a virulence plasmid (Mainil J.G. *et al*., 2000).

*1.4.2 Diarrhoeagenic E. coli (DEC)*

Diarrhoeagenic *E. coli* (DEC) is the *E. coli* pathogroup with the highest variability, involving both the virulence mechanisms and the complexity of the genomes. As a matter of fact, DEC pathogroup is subdivided in six different pathotypes: Enteropathogenic *E. coli* (EPEC), Verocytotoxin-producing *E. coli* (VTEC), Enterotoxigenic *E. coli*, Enteroaggregative *E. coli* (EAEC), Enteroinvasive *E. coli* (EIEC) and Diffusely-Adherent *E. coli* (DAEC).

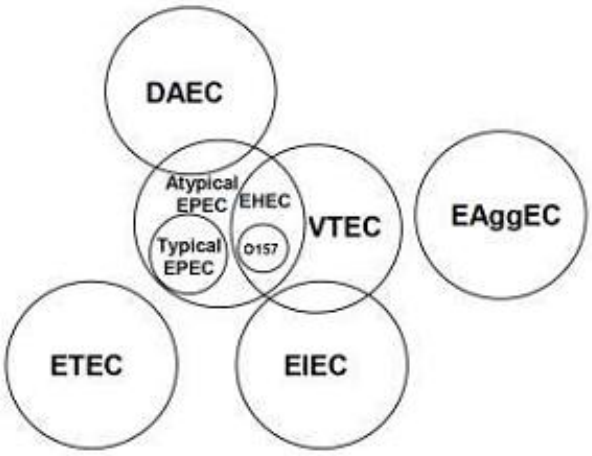



*Figure 2. Schematic representation of phylogenetic relations among the different pathotypes of Diarroheagenic E. coli (Adapted from: Donnenberg M. (Ed.) Academic Press 2002.)*

The phylogenetic relations existing between the different DEC pathotypes are particularly complex. Strains belonging to VTEC, EPEC, EIEC and DAEC pathotypes share several genetic determinants belonging to the "persistent" set of genes in the pangenome (Figure 1). Some groups, such as EPEC and VTEC share even a greater number of mobile genetic elements vehiculating virulence factors including some of the "volatile" genes indicating a closer correlation between the pathotypes (Figure 2).

Phylogenetic evidences indicate that these two pathotypes could derive from a common ancestor or from convergent evolutionary events, representing a good model for understanding evolutionary relations which have driven the emergence of the existing pathotypes of *E. coli*.

A brief overview of DEC pathotypes is presented here, with particular attention for Verocytotoxin-producing *E. coli*, which will be the focus of the present piece of research.

Enteropathogenic *E. coli* (EPEC)

After the first association of *E. coli* with outbreaks of infantile diarrhoea in the United Kingdom in the 1940s (Bray J., 1945), similar outbreaks were observed in several Western countries and found to be associated with particular serotypes of *E. coli*, collectively referred to as enteropathogenic *E. coli* (EPEC).

EPEC was the first pathotype of *E. coli* to be described. Although large outbreaks of infant diarrhoea due to EPEC have now largely disappeared from industrialized countries, EPEC remain an important cause of potentially fatal infant diarrhoea in low-income countries. For decades, the mechanisms by which EPEC caused diarrhoea remained unknown. However, since 1979, numerous advances in the understanding of the pathogenesis of EPEC diarrhoea have been made, such that EPEC is now among the best characterised pathogenic *E. coli*. (Kaper J.B. *et al*. 2004).

The main virulence factors governing the adhesion mechanism of EPEC strains to the intestinal mucosa are vehiculated by a PAI termed *locus of enterocyte effacement* (LEE) (McDaniel T.K. *et al*., 1995), which governs the production of typical *attaching-and-effacing* (A/E) lesions in the ileum: the bacteria stimulate large-scale cytoskeletal changes in the epithelial cells, resulting in the effacement of the microvilli brush border and the formation of a pedestal which accommodates the bacterial cell (Nataro J.P. & Kaper J.B., 1998), causing its intimate attachment to the eukaryotic cell. Following its first identification in EPEC strains (McDaniel T.K. *et al*., 1995), the LEE was later on shown to be present in some subpopulations of VTEC strains (Karmali M.A. *et al*., 2003).

Typical EPEC strains contain a large plasmid termed EAF (EPEC Factor for Adherence) which carries the *bfp* locus, coding for bundle forming pilus, and the *per* gene, whose product regulates the expression of the LEE genes involved in the A/E adhesion (Donnenberg M.S. *et al*., 1992). The presence of this plasmid is associated to EPEC strains responsible for the majority of gastroenteritis cases in low-income countries. Some EPEC strains, termed atypical EPEC (aEPEC), are EAF-negative and possess a virulence plasmid resembling the one typically present in the most pathogenic VTEC serogroups (see below).

Enterotoxigenic *E. coli* (ETEC)

Enterotoxigenic *E. coli* (ETEC) are a major cause of infantile diarrhoea in the developing world and are one of the leading causes of diarrhoea in travellers visiting these areas of the globe. The first description of ETEC in human infections dates back to the early 60s, when it was reported that certain *E. coli* isolates from the stools of children with diarrhoea elicited fluid secretion in ligated rabbit intestinal loops (Taylor J. *et al*., 1961). Ten years later, DuPont *et al*. showed that ETEC strains were able to cause diarrhoea in adult volunteers (DuPont H.L. *et al*., 1971).

ETEC strains were first recognized as causes of diarrheal disease in piglets. Studies on ETEC in this species first elucidated the mechanisms of disease and led to the characterization of two plasmid-encoded enterotoxins: the heat-Stable and the heat-Lable Toxins (ST and LT respectively) (Levine M.M., 1987; Mainil J.G. *et al*., 1998; Gomez-Duarte O.G. *et al*., 1999). LTs are similar to the enterotoxins produced

by *Vibrio cholerae* (Sixma T.K. *et al*., 1993), bind to membrane gangliosides and, upon internalization, produce an osmotic stress causing the release of great amounts of water in the faeces. STs exert a different molecular mechanism but, similarly to the LTs, cause a massive release of fluids in the gut.

Enteroinvasive *E. coli* (EIEC)

Enteroinvasive *E. coli* (EIEC) strains were first shown to be capable of causing diarrhoea during the studies on volunteers conducted by DuPont *et al*. in 1971 (DuPont H.L. *et al*., 1971). EIEC can cause an invasive inflammatory colitis and occasionally dysentery, but in most cases elicit watery diarrhoea indistinguishable from that caused by other pathogenic *E. coli* (Nataro J.P. & Kaper J.B., 1998). Numerous studies have shown that EIEC strains are biochemically, genetically and pathogenetically closely related to *Shigella* spp. (Nataro J.P. & Kaper J.B., 1998) and taxonomically indistinguishable at the species level (Pupo G.M. *et al*., 2000). Nevertheless, a nomenclature distinction is still maintained owing to the clinical significance of *Shigella*.

In particular, Enteroinvasive *E. coli* strains share with *Shigella flexneri* the large (140 MDa) virulence plasmid pInv. This plasmid vehiculates the invasion-related genes, which encode a type III secretion system, secreted proteins involved in the invasion phenotype (Ipa) and the *Shigella enterotoxin 2* (Nataro J.P. *et al*., 1995).

Enteroaggregative *E. coli* (EAEC)

In the early 1990s, two other diarrhoeagenic *E. coli* pathogroups were defined as distinct from EPEC for the mechanism of adhesion to the surface of Hep-2 cultured cells: Enteroaggregative *E. coli* (EAEC), showing a characteristic ‘‘stacked-brick’’ pattern of aggregation (Nataro J.P. *et al*., 1987), and Diffusely adherent *E. coli* (DAEC).

Pathogenesis of EAEC infections is believed to be initiated with adherence to the terminal ileum and colon in an aggregative, stacked-brick-type pattern by means of one of several different hydrophobic aggregative adherence fimbriae (AAFs), whose genetic determinants are vehiculated by a 55–65 MDa plasmids, collectively called pAA (Vial P.A. *et al*., 1988). The same plasmid also encodes most of the other known virulence factors of EAEC: the gene encoding a small protein termed antiaggregation protein (Aap), or dispersin, which neutralizes the strong negative charge of the LPS helping the displaying of the positively-charged AAFs out from the surface (Harrington S.M. *et al*., 2006); the *aat* operon encoding the ABC transporter responsible for the secretion of dispersin (Nishi J. *et al*., 2003); the 108 kDa plasmid-encoded toxin (Pet) (Eslava C. *et al*., 1998); the EAST1 cytotoxin, homologous to the ST of ETEC; the AggR global transcriptional activator, which regulates the expression of all these virulence factors.

In addition to these plasmid-encoded genes, other virulence genes are vehiculated by chromosomally located PAIs: the *setBA* genes, encoding ShET1 cytotoxin, are contained in a the 117 kb PAI inserted at the *pheU*-tRNA locus (Harrington S.M. *et al*., 2006); another PAI, integrated in *glyU* locus, encodes a type III secretion; a third PAI vehiculates genes for a transcriptional regulator and type III effector proteins and is inserted in *selC* (Harrington S.M. *et al*., 2006).

Diffusely Adherent *E. coli* (DAEC)

Diffusely Adherent *E. coli* (DAEC) are defined by a diffuse pattern of adherence on Hep-2 cultured cells with little aggregation. DAEC strains are associated with watery diarrhoea that can become persistent in young children (Le Bouguenec C., 1999).

The term "diffusely adherent *E. coli*" was initially used to refer to any HEp-2-adherent *E. coli* strain that did not form EPEC-like microcolonies. After the discovery of EAEC, DAEC have started to be recognized as an independent category of potentially diarrhoeagenic *E. coli*. Bilge *et al*. have described a surface fimbria produced by these strains and designated F1845, responsible for the DA phenotype (Bilge S.S. *et al*., 1989), which is encoded by genes that can be located either on the chromosome or vehiculated by a plasmid (Nataro J.P. & Kaper J.B., 1998). Beside the fimbriae, other adhesins, known as the afimbrial or nonfimbrial adhesins, have been associated with an amorphous, outer membrane-associated structure on the surface of DAEC strains as well as in some ExPEC strains (Soto G.E. & Hultgren S.J., 1999). The first determinant to be identified encoding an adhesin not associated with visible fimbriae on the bacterial surface was the one specified by the *afa* operon, encoding factors that can be involved either in the adhesion or in the internalization in the host cells, as well as in the cell signalling on the surface of the intestinal epithelium promoting tight attachment of bacteria.

Verocytotoxin-producing *E. coli* (VTEC)

Verocytotoxin-producing *E. coli* (VTEC) is one of the most recently defined *E. coli* pathotypes.

The first evidence of VTEC infection in humans dates back to 1983 (Riley L.W. *et al*., 1983) when two outbreaks of a distinctive gastrointestinal illness characterized by severe bloody diarrhoea occurred in the US. In both the episodes stool cultures from patients yielded a previously rarely isolated *E. coli* serotype, O157:H7. The outbreaks were associated with the ingestion of undercooked hamburgers at a fast-food restaurant chain and the disease was designated haemorrhagic colitis (HC). Similarly, the strains belonging to this pathotype were initially termed Enterohaemorragic *E. coli* (EHEC).

In the same year, Karmali *et al*. (Karmali M.A. *et al*., 1983) reported the association of sporadic cases of haemolytic uremic syndrome (HUS) with a faecal cytotoxin, termed Verocytotoxin (VT), and with Verocytotoxin-producing *E. coli* in the stools of patients. HUS is characterised by the triad of symptoms including acute renal failure, thrombocytopenia and microangiopathic hemolytic anemia and it is usually preceded by a bloody diarrhoeal illness indistinguishable from HC. The production of the Verocytotoxin was then demonstrated in the *E. coli* O157:H7 from the human cases and a novel and increasingly important class of enteric pathogens was recognised: Verocytotoxin-producing *E. coli* (VTEC), or Shiga-toxin producing *E. coli* (STEC).

Many efforts have been made in the last three decades to unveil the virulence factors implied in the pathogenesis of the disease caused by VTEC strains. However the complete set of virulence genes, generally referred to as the "virulome", still needs to be completely identified.

Verocytotoxins are the main virulence feature of VTEC. The holotoxins are about 70 kDa and consist of a single catalytic 32 kDa subunit and a pentamer of small 7.7 kDa subunits involved in the binding to specific receptors on the surface of target cells (O'Brien A.D. & Holmes R.K., 1987). Following the binding, the VTs are internalized by a receptor-mediated endocytosis mechanism and reach the

endoplasmic reticulum where they are activated through a proteolytic cleavage of the catalytic subunit (Sandvig K. *et al*., 1992). The activated toxin exerts an N-glycosidicasic activity on the 28S rRNA, causing the protein synthesis inhibition (Paton J.C. & Paton A.W., 1998).

VTs include two main antigenically distinct types, VT1 and VT2, and numerous subtypes identified by differences in the DNA sequence of the coding genes. Three subtypes of VT1 and seven subtypes of VT2 have been described so far, which can be found in any combination of type and subtype into a single VTEC (Persson S. *et al*., 2007). The VT-coding genes (*vtx*) are carried by bacteriophages belonging to the lambda family (Reid S.D. *et al*., 2000), which are usually maintained in a lysogenic state in the bacterial chromosome but retain the capability to enter the lytic cycle and move from a host to the other spreading the *vtx*-genes (Plunkett G. *et al*., 1999).

The presence of the *vtx*-converting bacteriophages is necessary condition for *E. coli* strains to cause the disease, though it seems not to be sufficient. As a matter of fact, VTEC associated with HC or HUS produce additional virulence factors boosting their pathogenicity. Some of these VTEC strains possess the *Locus of Enterocyte Effacement* (LEE) PAI encoding factors responsible for the "attaching and effacing" (A/E) hystopathological lesion (McDaniel T.K. *et al*, 1995). As in EPEC strains, the LEE encodes a type III secretion system (T3SS), consisting of a needle used to directly inject bacterial effectors into the host cell, causing the rearrangement of the cytoskeleton of the enterocyte and the typical brush border effacement. In addition to the T3SS, the LEE encodes an outer membrane protein called intimin (*eaeA* gene), which mediates the direct binding of the bacterium to the host cell surface, and its translocated receptor, Tir, which is vehiculated into the host cell plasma membrane by the T3SS (Yoon J.W. & Hovde C.J., 2008).

The presence of the LEE locus marks a subpopulation of VTEC which has been frequently associated to the most severe forms of human infections: HC and HUS.

Several other virulence genes have been described in VTEC strains isolated from human cases of severe disease, all carried by mobile genetic elements.

The PAI known as OI-122 hosts a large gene, termed *efa1/lifA*, encoding an immunomodulator, which inhibits the host lymphocyte activation. This PAI is consistently present in LEE-positive VTEC strains but is also present in the EPEC strains, suggesting the interaction of factors encoded by the two PAIs in governing the induction of the A/E lesion upon infection. Interestingly, OI-122 and the LEE have been found to be physically associated in a mosaic PAI in some VTEC and EPEC strains, suggesting the possibility that the hybrid PAI may have been first acquired as a co-integrated structure that may have undergone recombination events leading to the separation on two PAIs in several VTEC clones (Morabito S. *et al*., 2003).

VTEC strains isolated from HC and HUS in humans also possess a large virulence plasmid, called pO157 in VTEC of serogroup O157, vehiculating a number of additional virulence factors including the enterohaemolysin (*ehxA* gene), a catalase peroxidase (*katP* gene) and a serine-protease (*espP* gene) (Brunder W. *et al*., 1996). The pO157 also carries *toxB*, whose product seems to have a function similar to that of Efa1/lifA (Makino K. *et al*., 1998; Deng W. *et al*., 2012).

Some VTEC do not have the LEE locus and the pO157 virulence plasmid but are isolated from human cases of infection. These strains generally do not cause severe disease although they have been rarely isolated from bloody diarrhoea and HUS.

Such LEE-negative VTEC usually have other virulence-associated genes complementing the action of the LEE locus in the colonization of the host gut. The LEE-negative VTEC strains belonging to O113 serogroup, as an example, carry a virulence plasmid enconding a number of unique virulence-associated determinants including the autotransporter protein EpeA (Leyton D.L. *et al*., 2003), the autoagglutinating adhesin Saa (Paton A.W. *et al*., 2001) and a subtilase-like serine protease cytotoxin, termed Subtilase cytotoxin (SubAB) (Paton A.W. *et al*., 2004).

**1.5 Did VTEC emerge as a toxigenic clone of "attaching and effacing *Escherichia coli*"?**

VTEC are a heterogeneous population including strains with different assets of virulome. An important group of VTEC causes severe disease in humans such as the Haemorrhagic Colitis and the Haemolytic Uremic Syndrome. This VTEC group shares with EPEC genomic and pathogenetic analogies. They both have the LEE locus and cause the attaching and effacing (A/E) lesion to the enterocyte as the main mechanism for the host colonization (Nataro J.P. & Kaper J.B., 1998). Due to this feature, the two groups have been frequently included in a super-group termed "Attaching-and-Effacing *Escherichia Coli*" (AEEC). The genomic similarity within the AEEC pathotype is not restricted to the LEE locus. As a matter of fact, other PAIs, probably involved in the A/E phenotype, are present in all the AEEC strains. At a certain moment in the evolution of AEEC, some strains evolved the capability to elaborate the verocytotoxins dividing this group into two distinct subpopulations marked by differences in the virulence features and the epidemiology of the infection: EPEC and VTEC.

Typical EPEC strains are responsible for prolonged diarrhoea in developing countries with inter-human transmission of the infection through the classical oral-faecal route. Conversely VTEC are zoonotic pathogens and the diffusion of the infections is mainly mediated by food of animal origin or by environmentally contaminated vehicles. The natural reservoir of VTEC is represented by the gastrointestinal tract of ruminants, particularly cattle (Nataro J.P. & Kaper J.B., 1998).

As with most of the food-borne bacterial pathogens, VTEC infections mainly occur in industrialized countries where they can cause very large outbreaks, being the spreading of cases facilitated by the large-capacity distribution of food items.

Some EPEC strains can be placed at the interface between typical EPEC and VTEC. These strains have been defined as a group during the Second International Conference on EPEC, held in São Paulo, Brazil in 1995, where the strains of *E. coli* possessing the LEE locus but lacking the VT genes and the EAF plasmid have been termed «atypical EPEC» (aEPEC). Studies on the characterization of aEPEC showed that these strains have a genomic asset resembling that of VTEC (Trabulsi L.R. *et al*., 2002). As a matter of fact, they usually possess a large virulence plasmid vehiculating the operon encoding the enterohaemolysin and other virulence genes also present in the pO157 plasmid (Cookson A.L. *et al*., 2007; Hernandes R.T. *et al*., 2009).

Several theories have been proposed for the emergence of aEPEC strains. They could represent VTEC strains that simply lost the ability to produce VTs following the excision of the prophage vehiculating *vtx*

genes. However, some evidences suggest that aEPEC emerged from tEPEC following a "plasmid displacement" event (Wick L.M. *et al*., 2005).

One of the typical features of the virulence plasmid pO157, is the presence of the large virulence gene *toxB*. This gene, in its entire form or as remnant, is quite diffuse in the virulence plasmids present in several VTEC serogroups and part of *toxB* gene is also present in the EAF plasmid of the prototype tEPEC strain E2348/69 (Tobe T. *et al*., 1999). Similarly to *toxB*, the *per* gene, located on the EAF plasmid and encoding a regulator of some LEE genes in tEPEC, has been detected in aEPEC, lacking the genes coding for the Bfp fimbriae, considered a hallmark for the EAF plasmid (Contreras C.A. *et al*., 2010).

These observations suggest that the two MGE may have come in contact and that a plasmid displacement event may have occurred causing the substitution of the EAF plasmid with a pO157-derived MGE.

The appearance of aEPEC may be considered as an intermediate step in the evolution of VTEC according to the following step-wise model: an *E. coli* strain may have acquired the LEE locus and the MGEs encoding the accessory adhesins (e.g. OI-122) and the non-LEE encoded effectors probably from different bacterial species, thus triggering the selection of an *attaching-and-effacing E. coli* ancestor. Afterwards, the acquisition of a plasmid carrying the *bfp* operon could have occurred into the chimeric *E. coli*, originating a tEPEC-like ancestor. Finally, the displacement of the *bfp*-carrying plasmid (EAF) by a pO157-like structure, followed by the acquisition of a *vtx*-converting bacteriophage, may have led to the emergence of the VTEC prototype.

Such a model has been proposed for the evolution of VTEC O157 from a tEPEC of serotype O55:H7 (Figure 3) (Wick L.M. *et al*., 2005).

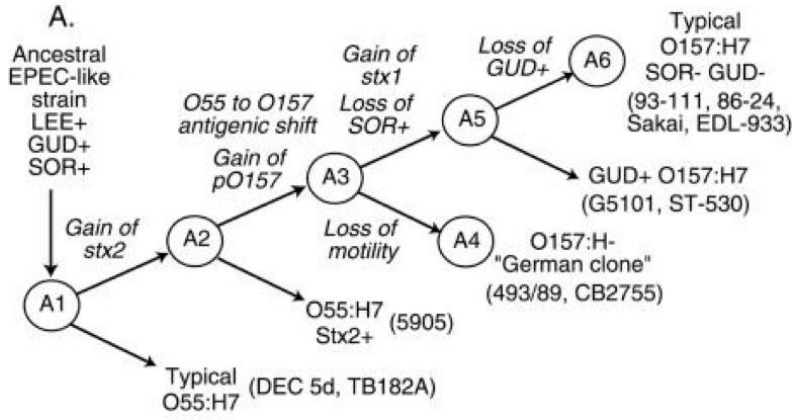



*Figure 3. Model for the evolution of VTEC O157:H7 from a tEPEC belonging to serotype O55:H7 (Wick LM et al., 2005)*

**1.6 Emergence of new pathogroups: the O104 example**

Horizontal gene transfer can cause pathogenic bacterial populations to emerge from harmless species. Moreover the exchange of genetic material by HGT may favour the formation of new pathotypes by mixing virulence features deriving from different pathogenic types in a single bacterial cell.

The appearance of the VTEC O104:H4 *E. coli* strain, which caused the large outbreak of infections in Germany in May 2011, is paradigmatic of this scenario. The episode was one of the largest outbreaks

ever occurred, it counted about 4000 ill and more than 850 HUS cases with the heavy toll of 50 deaths (Frank C. *et al*., 2011; Jansen A. & Kielstein J.T., 2011).

The causative agent did not match the typical definition of pathogenic VTEC. It did not possess the LEE locus or the enterohaemolysin-coding gene and it did not belong to any of the serogroups known to cause HUS. The whole-genome sequencing of the outbreak strain revealed a rare asset of virulence determinants where the bacteriophage carrying the VT-coding genes was inserted in the genomic backbone of a different *E. coli* pathotype: the Enteroaggregative *E. coli* (EAEC).

This particular genomic arrangement was not novel. It had been sporadically observed during the last twenty years. In the early 1990s a VT2-producing O111 EAEC strain was isolated during a small outbreak of HUS in France (Morabito S. *et al*., 1999). An EAEC that produced VT2 has been isolated from cases of HC and HUS in patients with HIV in Bangui, Central African Republic (Mossoro C. *et al*., 2002) and a few sporadic cases have been reported in the EU (Scheutz F. *et al*., 2011; Mellmann A. *et al*., 2008). In 2009 a single case of infection, described in Italy, was later on identified as caused by a VT-producing EAEC O104:H4 showing a molecular profile similar to the strains that caused the German outbreak (This study, Scavia G. *et al*., 2011). Finally, an HUS case ascribed to infection with VT-producing EAEC O104:H4 has been reported in Finland in 2010 (Scheutz F. *et al*., 2011).

It has to be considered that for such a chimeric pathogen to emerge, an EAEC strain should have come in contact with a VT-converting bacteriophage. In general this scenario should not occur easily. In fact EAEC have a human reservoir, circulate predominantly in the human host and the infections are diffuse mainly in developing countries. Conversely, VTEC have an animal reservoir and VTEC infections are common in industrialized countries and thus the two pathotypes are kept separated epidemiologically and geographically.

Nevertheless, EAEC can easily survive in the environment where they can form strong biofilms. In developing countries, the burden of infection caused by these microorganisms is very high. In these countries, they are massively released with the faeces in the sewage and treatment of wastewaters are generally not in place or are scarcely effective, causing the contamination of the aquatic environment. Water in turn represents the main re-infection vehicle for humans leading to the establishment of an amplification of the pathogen's populations.

*Figure 4. Promiscuous conditions illustrating the overlapping between the human and the animal reservoir of pathogenic E. coli in low-income countries (Pakistan Associated Press*/Pervez Masih)

In order to get a figure of this phenomenon, it has been estimated that about 2.6 billion of people in underdeveloped areas of the globe do not have access to water suitable for potable or personal hygiene uses (http://www.unwater.org/statistics_san.html). Moreover, it is not infrequent that, in these parts of the world, the same aquatic environment is shared by human and animals, including ruminants, which are the main reservoir of VTEC, setting the proper condition for the contacts between EAEC and the VT-converting bacteriophages to occur (Figure 4).

The coexistence of human and animal *E. coli* strains in the same ecological niche could thus have triggered the remodelling of the versatile genome characterizing this bacterial species and may have facilitated the exchange of mobile genetic elements through events of horizontal gene transfer.

## 1.7 References


- Bergthorsson U, Ochman H. 1995. **Heterogeneity of genome sizes among natural isolates of *Escherichia coli***. J Bacteriol. 177(20):5784-9.
- Bergthorsson U, Ochman H. 1998. **Distribution of chromosome length variation in natural isolates of *Escherichia coli***. Mol Biol Evol 15:6–16.
- Bern C. 2004. **Diarrhoeal Diseases In The Global Epidemiology of Infectious Diseases.** ed. Murray C.J.L., Lopez A.D. and Mathers C.D. 1-27. Vol. 4 of Global Burden of Disease and Injury Series. Geneva: World Health Organization.
- Bilge SS, Clausen CR, Lau W, Moseley SL. 1989. **Molecular characterization of a fimbrial adhesin, F1845, mediating diffuse adherence of diarrhea-associated *Escherichia coli* to HEp-2 cells.** J. Bacteriol. 171: 4281–4289.
- Blattner FR, Plunkett G 3rd, Bloch CA *et al*. 1997. **The complete genome sequence of *Escherichia coli* K-12**. Science. 277(5331):1453-62.
- Boyd EF, Brussow H. 2002. **Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved**. Trends in Microbiology 10, 521–529.
- Bray J. 1945. **Isolation of antigenically homogeneous strains of *Bact. coli neopolitanum* from summer diarrhoea of infants**. J. Pathol. Bacteriol. 57, 239–247.
- Brunder W, Schmidt H, Karch H. 1996. **KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic *Escherichia coli* O157:H7.** Microbiology. 142 ( Pt 11):3305-15.
- Campbell A, 1996. **Bacteriophages**. In: Neidhardt F, Curtis III J, Ingraham, J. (Eds.) *Escherichia coli* and *Salmonella*:Cellular and Molecular Biology. ASM Press, Washington, DC, pp. 2325–2338.
- Caprioli A, Falbo V, Ruggeri FM *et al*. 1987. **Cytotoxic necrotizing factor production by hemolytic strains of *Escherichia coli* causing extraintestinal infections.** J Clin Microbiol. 25(1):146-9.
- Casjens S. 2003. **Prophages and bacterial genomics: what have we learned so far?** Mol. Microbiol. 49: 277-300.
- Contreras CA, Ochoa TJ, Lacher DW *et al*. 2010. **Allelic variability of critical virulence genes (*eae, bfpA* and *perA*) in typical and atypical enteropathogenic *Escherichia coli* in Peruvian children.** J Med Microbiol. 59(Pt 1):25-31.
- Cookson AL, Bennett J, Thomson-Carter F, Attwood GT. 2007. **Molecular subtyping and genetic analysis of the enterohemolysin gene (*ehxA*) from Shiga toxin-producing *Escherichia coli* and atypical enteropathogenic *E. coli*.** Appl Environ Microbiol. 73(20):6360-9.
- Deng W, Yu HB, de Hoog CL *et al*. 2012. **Quantitative proteomic analysis of type III secretome of enteropathogenic *Escherichia coli* reveals an expanded effector repertoire for attaching/effacing bacterial pathogens**. Mol Cell Proteomics. 11(9):692-709.

- Dobrindt U, Chowdary MG, Krumbholz G, Hacker J. 2010. **Genome dynamics and its impact on evolution of *Escherichia coli***. Med Microbiol Immunol 199:145–154.
- Dobrindt U, Hochhut B, Hentschel U *et al*. 2004. **Genomic islands in pathogenic and environmental microorganisms**. Nat Rev Microbiol 2:414–424.
- Dobrindt U. **(Patho-)Genomics of *Escherichia coli*. 2005**. International Journal of Medical Microbiology 295. 357–371.
- Donnenberg M. (Ed.) 2002. ***Escherichia Coli*: Virulence Mechanisms of a Versatile Pathogen.** Academic Press.
- Donnenberg MS, Girón JA, Nataro JP, Kaper JB. 1992. **A plasmid-encoded type IV fimbrial gene of enteropathogenic *Escherichia coli* associated with localized adherence**. Mol Microbiol. 6(22):3427-37.
- DuPont HL, Formal SB, Hornick RB *et al*. 1971. **Pathogenesis of *Escherichia coli* diarrhea.** N. Engl. J. Med. 285:1–9.
- Eslava C, Navarro-García F, Czeczulin JR *et al*. 1998. **Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*.** Infect Immun. 66(7):3155-63.
- Frank C, Werber D, Cramer JP *et al*. HUS Investigation Team. 2011. **Epidemic profile of Shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany.** N Engl J Med. 365(19):1771-80.
- Freeman VJ. 1951. **Studies on the virulence of bacteriophage-infected strains of *Corynebacteria diphtheriae***. J. Bacteriol. 61: 675-688.
- Gomez-Duarte OG, Ruiz-Tagle A, Gomez DC *et al*. 1999. **Identification of *lngA*, the structural gene of longus type IV pilus of enterotoxigenic *Escherichia coli*.** Microbiology 145, 1809–1816.
- Gomez-Lus, R, 1998. **Evolution of bacterial resistance to antibiotics during the last three decades**. Int. Microbiol. 1, 279–284.
- Grozdanov L, Raasch C, Schulze J *et al*. 2004. **Analysis of the genome structure of the nonpathogenic probiotic *Escherichia coli* strain Nissle 1917**. J Bacteriol. 186(16):5432-41.
- Guyer, DM, Kao JS, Mobley HL. 1998. **Genomic analysis of a pathogenicity island in uropathogenic *Escherichia coli* CFT073: distribution of homologous sequences among isolates from patients with pyelonephritis, cystitis, and catheter-associated bacteriuria and from fecal samples**. Infect. Immun. 66, 4411–4417.
- Hacker J, Carniel E. 2001. **Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes**. EMBO Rep. 2, 376–381.
- Hacker J, Kaper JB. 2002. **Pathogenicity islands and the evolution of pathogenic microbes**. Curr. Top. Microbiol. Immunol. 264 (1,2).
- Harrington SM, Dudley EG, Nataro JP. 2006. **Pathogenesis of enteroaggregative *Escherichia coli* infection.** FEMS Microbiol Lett. 254(1):12-8.
- Hayashi T, Makino K, Ohnishi M *et al*. 2001. **Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12**. DNA Res. 8(1):11-22. Erratum in: DNA Res. 8(2):96.

- Hernandes RT, Elias WP, Vieira MA, Gomes TA. 2009. **An overview of atypical enteropathogenic *Escherichia coli*.** FEMS Microbiol Lett. 297(2):137-49.
- Jansen A, Kielstein JT. 2011. **The new face of enterohaemorrhagic *Escherichia coli* infections.** Euro Surveill. 16(25).
- Johnson JR, O'Bryan TT, Kuskowski M, Maslow JN. 2001. **Ongoing horizontal and vertical transmission of virulence genes and *papA* alleles among *Escherichia coli* blood isolates from patients with diverse-source bacteremia.** Infect. Immun. 69, 5363–5374.
- Johnson JR, Russo TA. 2005. **Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli***. Int J Med Microbiol. 295(6-7):383-404.
- Kaper JB, Nataro JP, Mobley HL. 2004. **Pathogenic *Escherichia coli*.** Nat Rev Microbiol. 2(2):123-40.
- Karmali MA, Mascarenhas M, Shen S *et al*.. 2003. **Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease**. J Clin Microbiol. 41(11):4930-40.
- Karmali MA, Steele BT, Petric M, Lim C. 1983. **Sporadic cases of haemolytic uremic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools.** Lancet i:619–620.
- Kelly BG, Vespermann A, Bolton DJ. 2009. **The role of horizontal gene transfer in the evolution of selected foodborne bacterial pathogens**. Food and Chemical Toxicology 47. 951–968.
- Köhler CD, Dobrindt U. 2011. **What defines extraintestinal pathogenic *Escherichia coli*?** Int J Med Microbiol. 301(8):642-7.
- Kosek M, Bern C, Guerrant RL. 2003. **The global burden of diarrhoeal disease, as estimated from studies published between 1992–2000**. Bull. World Health Organ. 81: 197–204.
- Le Bouguenec C. 1999. **Diarrhea-associated diffusely adherent *Escherichia coli*.** Clin Microbiol Rev. 12: 180–181.
- Levine MM. 1987. ***Escherichia coli* that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent.** J Infect Dis. 155(3):377-89.
- Leyton DL, Sloan J, Hill RE *et al*. 2003. **Transfer region of pO113 from enterohemorrhagic *Escherichia coli*: similarity with R64 and identification of a novel plasmid-encoded autotransporter, EpeA.** Infect Immun. 71(11):6307-19.
- Mainil JG, Daube G, Jacquemin E *et al*. 1998. **Virulence plasmids of enterotoxigenic *Escherichia coli* isolates from piglets.** Vet. Microbiol. 62, 291–301.
- Mainil JG, Gérardin J, Jacquemin E. 2000. **Identification of the F17 fimbrial subunit- and adhesin-encoding (*f17A* and *f17G*) gene variants in necrotoxigenic *Escherichia coli* from cattle, pigs and humans**. Vet Microbiol. 73(4):327-35.
- Makino K, Ishii K, Yasunaga T *et al*. 1998. **Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic *Escherichia coli* O157:H7 derived from Sakai outbreak**. DNA Research 5(1), 1–9.

- McDaniel TK, Jarvis KG, Donnenberg MS, Kaper, JB. 1995. **A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens**. Proc. Natl. Acad. Sci. USA 92, 1664–1668.
- Mellmann A, Bielaszewska M, Kock R *et al*. 2008. **Analysis of collection of hemolytic uremic syndrome-associated enterohemorrhagic *Escherichia coli*.** Emerg Infect Dis. 14:1287–1290.
- Morabito S, Karch H, Schmidt H *et al*. 1999. **Molecular characterisation of verocytotoxin-producing *Escherichia coli* of serogroup O111 from different countries.** J Med Microbiol. 48(10):891-6.
- Morabito S, Tozzoli R, Oswald E, Caprioli A. 2003. **A mosaic pathogenicity island made up of the locus of enterocyte effacement and a pathogenicity island of *Escherichia coli* O157:H7 is frequently present in attaching and effacing *E. coli*.** Infect Immun. 71(6):3343-8.
- Mossoro C, Glaziou P, Yassibanda S *et al*. 2002. **Chronic diarrhea, hemorrhagic colitis, and hemolytic-uremic syndrome associated with HEp-2 adherent *Escherichia coli* in adults infected with human immunodeficiency virus in Bangui, Central African Republic.** J Clin Microbiol. 40(8):3086-8.
- Nataro JP, Kaper JB, Browne RR *et al*. 1987. **Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells.** Pediatr. Infect. Dis. J. 6:829–831.
- Nataro JP, Kaper JB. 1998. **Diarrheagenic *Escherichia coli***. Clin. Microbiol. Rev. 11, 142–201.
- Nataro JP, Seriwatana J, Fasano A *et al*. 1995. **Identification and cloning of a novel plasmid-encoded enterotoxin of enteroinvasive *Escherichia coli* and *Shigella* strains.** Infect Immun. 63(12):4721-8.
- Nishi J, Sheikh J, Mizuguchi K *et al*. 2003. **The export of coat protein from enteroaggregative *Escherichia coli* by a specific ATP-binding cassette transporter system.** J Biol Chem. 278(46):45680-9.
- O'Brien AD, Holmes RK. **Shiga and Shiga-like toxins.** 1987. Microbiol. Rev., 51:206–220.
- Ochman H, Lawrence JG, Groisman EA. 2000. **Lateral gene transfer and the nature of bacterial innovation**. Nature 405:299–304.
- Paton AW, Srimanote P, Talbot UM *et al*. 2004. **A new family of potent AB(5) cytotoxins produced by Shiga toxigenic *Escherichia coli*.** J Exp Med. 200:35–46.
- Paton AW, Srimanote P, Woodrow MC, Paton JC. 2001. **Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans.** Infect Immun. 69(11):6999-7009.
- Paton JC, Paton AW. 1998. **Pathogenesis and diagnosis of Shiga Toxin-producing *Escherichia coli* infections.** Clin. Microbiol. Rev. 11(3):450–479.
- Perna NT, Plunkett G 3rd, Burland V *et al*. 2001. **Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7**. Nature. 409: 529-533.
- Persson S, Olsen KE, Ethelberg S, Scheutz, F. 2007. **Subtyping method for *Escherichia coli* shiga toxin (verocytotoxin) 2 variants and correlation to clinical manifestations.** J Clin Microbiol. 45:2020-45.

- Plunkett G, Rose DJ, Durfee TJ, Blattner FR. 1999. **Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product**. J. Bacteriol. 181:1767-1778.
- Pupo GM, Lan R, Reeves PR. 2000. **Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics.** Proc. Natl Acad. Sci. USA 97, 10567–10572.
- Rasko DA, Rosovitz MJ, Myers GS *et al*. 2008. **The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates**. J Bacteriol 190: 6881–6893.
- Redford P, Welch RA. 2002. **Extraintestinal *Escherichia coli* as a model system for the study of pathogenicity islands**. Curr. Top. Microbiol. Immunol. 264, 15–30.
- Reid SD, Herbelin CJ, Bumbaugh AC *et al*. 2000. **Parallel evolution of virulence in pathogenic *Escherichia coli***. Nature. 406:64-67.
- Riley LW, Remis RS, Helgerson SD *et al*. 1983. **Hemorrhagic colitis associated with a rare *Escherichia coli* serotype**. N Engl J Med. 308(12):681-5.
- Sandvig K, Garred O, Prydz K *et al*. 1992. **Retrograde transport of endocytosed Shiga toxin to the endoplasmic reticulum.** Nature. 358:510–512.
- Scavia G, Morabito S, Tozzoli R *et al*. 2011. **Similarity of Shiga toxin-producing *Escherichia coli* O104:H4 strains from Italy and Germany**. Emerg Infect Dis. 17(10):1957-8.
- Scheutz F, Nielsen EM, Frimodt-Møller J *et al*. 2011. **Characteristics of the enteroaggregative Shiga toxin/verotoxin-producing *Escherichia coli* O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011.** Euro Surveill. 16(24).
- Schmidt H, Hensel M. 2004. **Pathogenicity islands in bacterial pathogenesis.** Clin Microbiol Rev. 17(1):14-56.
- Sixma TK, Kalk KH, van Zanten BA *et al*. 1993. **Refined structure of *Escherichia coli* heat-labile enterotoxin, a close relative of cholera toxin.** J Mol Biol. 230(3):890-918.
- Smith JL, Fratamico PM, Gunther NW. 2007. **Extraintestinal pathogenic *Escherichia coli***. Foodborne Pathog Dis. 4(2):134-63.
- Soto GE, Hultgren SJ. 1999. **Bacterial adhesins: common themes and variations in architecture and assembly.** J Bacteriol. 181(4):1059-71.
- Taylor J, Wilkins MP, Payne JM. 1961. **Relation of rabbit gut reaction to enteropathogenic *Escherichia coli*.** Br. J. Exp. Pathol. 42:43–52.
- Tinsley CR, Bille E, Nassif X. 2006. **Bacteriophages and pathogenicity: more than just providing a toxin?** Microbes Infect. 8: 1365-1371.
- Tobe T, Hayashi T, Han CG *et al*. 1999. **Complete DNA sequence and structural analysis of the enteropathogenic *Escherichia coli* adherence factor plasmid.** Infect Immun. 67(10):5455-62.
- Touchon M, Hoede C, Tenaillon O *et al*. 2009. **Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths**. PLoS Genet 5:e1000344.

- Trabulsi LR, Keller R, Tardelli Gomes TA. 2002. **Typical and atypical enteropathogenic *Escherichia coli*.** Emerg Infect Dis. 8(5):508-13.
- Van Elsas JD, Semenov AV, Costa R, Trevors JT. 2011. **Survival of *Escherichia coli* in the environment:fundamental and public health aspects**. The ISME Journal. 5; 173–183.
- Vial PA, Robins-Browne R, Lior H *et al*. 1988. **Characterization of enteroadherent-aggregative *Escherichia coli*, a putative agent of diarrheal disease.** J. Infect. Dis. 158, 70–79.
- Welch RA, Burland, V, Plunkett G 3rd *et al*. 2002. **Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*.** Proc. Natl. Acad. Sci. USA 99, 17020–17024.
- Wick LM, Qi W, Lacher DW, Whittam TS. 2005. **Evolution of genomic content in the stepwise emergence of *Escherichia coli* O157:H7.** J Bacteriol. 187(5):1783-91.
- Yoon JW, Hovde CJ. 2008. **All blood, No stool: enterohemorrhagic *Escherichia coli* O157:H7 infection**. J. Vet. Sci. 9(3):219-231.

# 2.Aims of the work

# 2. AIMS OF THE WORK

The scope of the present work was to investigate the molecular bases of VTEC pathogenicity by determining their whole virulence genes asset. As a matter of fact VTEC are a heterogeneous pathotype of *E. coli* characterised by a huge variability in both the size and content of the virulome. This hindered the comprehension of its real extent and the identification of all the virulence genes really involved in the pathogenetic process of the VTEC-induced disease.

The knowledge of the virulence machinery of VTEC might also have reflections on contexts other than the pathogenesis of the infections. As an example, VTEC-infected patients can't be treated with antibiotic drugs, since their use has been proved to increase the production of the verocytotoxins by inducing an SOS-response in the bacterium. Furthermore a wide range of symptoms including diarrhoea and haemorrhagic diarrhoea characterizes the illness caused by VTEC. The latter is often prodromal to the development of the most severe manifestation of the infection, the HUS. This complex symptomatological picture accounts for a significant under-diagnosis of these infections and their frequent misdiagnosis with the subsequent use of invasive and inappropriate surgical intervention. In addition, VTEC are indistinguishable from other *E. coli* including those that are commensals and generally non-pathogenic. Therefore, the identification of appropriate pathogenicity flags, such as a definite set of virulence genes, may allow the development of suitable tools for a prompt diagnosis limiting the misuse of surgery. Moreover a timely recognition of a VTEC infection in a patient with diarrhoea would avoid the administration of antimicrobials reducing the risk of a progression towards the most severe form, the HUS.

As of today, the control of the infections is mainly based on prevention. Such a strategy includes the control of the vehicles of infection, mainly foodstuff, as well as the control of the animal reservoir.

The results of the proposed research can also improve such a preventive control strategy. Given the difficulties in selectively growing pathogenic VTEC *vs* the other commensal *E. coli* and the complexity of the available molecular assays aiming at detecting a high number of targets, the identification of virulence predictors could be of great benefit for the development of tools to be used for the detection of pathogenic VTEC strains in food matrices constituting the main vehicles of infections.

The limitations of the available detection strategies for VTEC in food, involving the detection of *vtx* and *eae* gene as well as the genes associated to a restricted number of serogroups, became evident during the large outbreak of VTEC O104:H4 infection occurred in Germany in May-June 2011. The characterization of the outbreak strain outlined the possibility of the emergence of new pathotypes of *E. coli* through the horizontal transfer of mobile genetic elements vehiculating factors involved in the pathogenic process. As a matter of fact the acquisition of a *vtx*-phage by an Enteroaggregative *E. coli* generated one of the most aggressive VTEC ever described.

A deeper knowledge of the complete asset of virulence genes composing the "virulome" of VTEC and of the MGEs involved in their dissemination would contribute to the preparedness necessary to eventually face infections from new arising combinations of virulence genes from different *E. coli* pathotypes, as occurred with the VT-producing EAEC O104:H4.

The general experimental strategy adopted in this piece of research has been the analysis of the genomic content of VTEC by using high throughput methodologies and paying particular attention to the mobile genetic elements due to their central role in the dissemination of virulence genes in *E. coli*.
The workplan has been divided in the following work packages, listed here along with the corresponding detailed objectives:

1. Characterization of the genetic determinants encoding a novel allelic variant of the Subtilase cytotoxin (SubAB):
   - Identification and characterization of the genetic determinant encoding the new SubAB variant ($SubAB_2$)
   - Functional analysis of $SubAB_2$
   - Identification and characterization of the mobile genetic element vehiculating $SubAB_2$
   - Study of the distribution of SubAB-coding genes in VTEC strains deriving from human and animal sources

   *Two publications presented in chapter 3*

2. Identification of a virulence-associated MGE of VTEC serogroups associated with severe human disease (OI-57):
   - Identification and characterization of genomic islands strongly associated with the most pathogenic VTEC serogroups
   - Study of the distribution of the MGE OI-57 in different *E. coli* pathotypes
   - Identification of new putative virulence predictors for VTEC strains belonging to the most pathogenic serogroups

   *Publication presented in chapter 4*

3. Analysis of VTEC strains possessing rare combinations of MGEs vehiculating virulence factors:
   - Analysis of the virulence determinants of an O104 VTEC strain isolated in Italy in 2009
   - Identification of the phylogenetic relationship with the O104 VT-producing Enteroaggregative *E. coli* strain that caused the outbreak of HUS in Germany in 2011
   - Development of molecular tools for the specific detection of the strain responsible for the German outbreak in food matrices.

   *Two publications presented in chapter 5*

# 3. Characterization of the genetic determinant of a novel allelic variant of the Subtilase cytotoxin

# 3. Characterization of the genetic determinant of a novel allelic variant of the Subtilase cytotoxin (SubAB)

*Introduction*

Subtilase (SubAB) is an $AB_5$ cytotoxin elaborated by some VTEC strains usually lacking the LEE locus. The production of SubAB has been first described in a VTEC O113 strain isolated from a human case of HUS in Australia in 2003 and later on identified in other VTEC serogroups. The action of SubAB has been proved to cause a stress in the endoplasmic reticulum of the host cell, inducing the unfolded protein response and finally inducing apoptosis. SubAB has been supposed to contribute to the pathogenetic process through a synergistic action with the VTs. In the prototype strain, SubAB-encoding operon is located on the large virulence plasmid pO113, also vehiculating *saa* gene, encoding an autoagglutinating adhesin.

*Results*

The production of SubAB by two VT-negative *E. coli* strains (ED 32 and ED 591) is reported for the first time. These strains had been isolated from unrelated cases of childhood diarrhoea. The characterisation of the SubAB-encoding genes showed a 90% nucleotidic sequence similarity with that of the prototype *subAB* vehiculated by pO113 plasmid (termed $subAB_1$), representing a new allelic variant of *subAB* (termed $subAB_2$). The two strains were also negative for the *saa* gene and harboured *subAB* operon on the chromosome, in association with *tia* gene, encoding an invasion determinant described in Enterotoxigenic *E. coli* strains associated to human disease. (Publication n.1: *Journal of Clinical Microbiology*, 48(1):178-183. Impact Factor 4.153).

The entire nucleotidic sequence of the chromosomal locus containing the $subAB_2$ operon in the ED 32 strain was determined, consisting in an 8 kb-long pathogenicity island that has been termed Subtilase-Encoding PAI (SE-PAI). The PAI was integrated in the *pheV*-tRNA locus and, besides *subAB* and *tia* genes, also vehiculates a gene encoding an integrase and a gene showing great sequence similarity with *shiA* gene of *Shigella*. Moreover, the distribution of the two *subAB* allelic variants has been investigated in *E. coli* strains belonging to different pathotypes and in a large collection of LEE-negative VTEC of human and ovine origin. The results confirmed that *subAB* genes are carried predominantly by VTEC, showing their presence in 72% and 86% of the LEE-negative strains from human cases of diarrhoea and from healthy sheep examined, respectively. Most of the *subAB*-positive strains identified (98%) possessed the $subAB_2$ allelic variant and were also positive for *tia*, suggesting the presence of the entire SE-PAI. (Publication n.2: *Clinical Microbiology and Infection*, Accepted in December 2012, see Appendix III. Impact Factor 4.54).

*Conclusions*

A new allelic variant of Subtilase cytotoxin ($SubAB_2$) has been identified and the distribution of the two variants has been analysed in a large panel of LEE-negative VTEC strains isolated from human and ovine

sources. The results showed that *subAB* is present in the vast majority of LEE-negative VTEC strains isolated from ovine sources and from human cases of disease and that *subAB*$_2$ is the prevalent allele in LEE-negative VTEC circulating in European countries. Altogether these findings indicate that sheep may represent an important reservoir for human infections with these strains and that *subAB* could play an important role in the pathogenesis of the disease. Moreover, the mobile genetic element vehiculating *subAB*$_2$ variant has been characterised, consisting in a PAI inserted in *pheV*-tRNA locus, which represents a hotspot of integration for genomic islands in *E. coli* genome. The use of this integration site by SE-PAI could provide an explanation for the mutual exclusion of this PAI and LEE locus, often integrated in the same chromosomal site. The SE-PAI resulted to harbour also genes which could play a role in the pathogenesis: beside *subAB*, the PAI vehiculates *shiA* gene, whose product is involved in the attenuation of the inflammatory response induced in the host upon *Shigella* infection and could exert a similar immunomodulator effect in *E. coli*; and *tia* gene, which encodes an invasion factor in ETEC strains and could play a similar role in the colonization process by different *E. coli* pathotypes. These observations suggest that LEE-negative VTEC strains express accessory virulence factors that could play an important role in the pathogenesis, acting with mechanisms alternative to the ability to cause *attaching-and-effacing* lesions, thus supplying the absence of the LEE locus.

JOURNAL OF CLINICAL MICROBIOLOGY, Jan. 2010, p. 178–183 Vol. 48, No. 1
0095-1137/10/$12.00 doi:10.1128/JCM.01648-09


# Production of the Subtilase $AB_5$ Cytotoxin by Shiga Toxin-Negative *Escherichia coli*[▽]

Rosangela Tozzoli, Alfredo Caprioli, Stefano Cappannella, Valeria Michelacci, Maria Luisa Marziano, and Stefano Morabito*

*Community Reference Laboratory for Escherichia coli, Istituto Superiore di Sanità, Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare, Rome, Italy*

Received 25 August 2009/Returned for modification 12 October 2009/Accepted 12 November 2009

**The subtilase cytotoxin (SubAB) is an $AB_5$ toxin described in certain Shiga toxin (Stx)-producing *Escherichia coli* (STEC) strains that usually lack the locus for enterocyte effacement (LEE). We report for the first time the production of SubAB by two Stx-negative *E. coli* strains, isolated from unrelated cases of childhood diarrhea. The characterization of the SubAB-coding genes showed a 90% nucleotide sequence similarity with that of the prototype *subAB*, located on the virulence plasmid of the STEC O113 strain 98NK2 (pO113). In both strains, *subAB* was physically associated with *tia*, an invasion genetic determinant of enterotoxigenic *E. coli*. The strains were negative for the *saa* gene, encoding an adhesin located on pO113 and present in many of the SubAB-positive strains described so far. PCR screening of 61 STEC and 100 Stx-negative *E. coli* strains in our collection revealed the presence of *subAB* in five LEE-negative STEC strains but not in the Stx-negative strains. *subAB* was contiguous to *tia* in three of the positive strains, which were all negative for *saa*. These results indicate that SubAB production is not restricted to STEC and suggest that a *subAB-tia* putative pathogenicity island is involved in the dissemination of *subAB* genes, as an alternative to plasmid pO113.**

Diarrheagenic *Escherichia coli* may produce $AB_5$ toxins constituted by an A subunit containing the catalytic activity and a pentamer of B subunits, usually involved in the binding to a cellular receptor. *E. coli* $AB_5$ toxins directly involved in the induction of illness include Shiga toxins (Stx) (28) and the heat-labile enterotoxins, related to cholera toxin (27). The subtilase cytotoxin (SubAB) is the prototype of a recently discovered $AB_5$ cytotoxin family produced by *E. coli* strains associated with human disease and that also produce Stx (19).

SubAB is constituted by a 35-kDa A subunit showing a subtilase-like serine protease activity and by five B subunits, related to a putative exported protein from *Yersinia pestis* and forming a pentamer which mediates binding to glycan receptors terminating in α-2-3-linked *N*-glycolylneuraminic acid on the target cell surface (2). This toxin was first described in the Stx-producing *E. coli* (STEC) strain 98NK2, associated with a small outbreak of hemolytic uremic syndrome (HUS) in South Australia (19). This strain belonged to serotype O113:H21 and did not possess the locus for enterocyte effacement (LEE), which governs the attaching-effacing mechanism of intestinal adhesion and represents a common feature of STEC strains associated with severe human disease (7).

In the 98NK2 strain, SubAB is encoded by two closely linked, cotranscribed genes (*subA* and *subB* [*subAB*]) located on a large, conjugative virulence plasmid designated pO113. This plasmid is also characterized by the presence of *saa* (20), a locus encoding the production of an autoagglutinating adhesin whose role in the colonization of the host intestinal mucosa has been hypothesized (20).

Since their first description in 2004, SubAB-coding genes have been identified in other LEE-negative STEC strains (4, 18), and it has been hypothesized that the toxin may contribute to the pathogenesis of the disease induced by these strains by a synergistic action with Stx (18, 19). SubAB causes a cytopathic effect (CPE) on Vero cells (19), and upon intraperitoneal injection in mice, it induces the typical features of the Stx-associated HUS, including extensive microvascular damage, thrombosis, and necrosis in brain, kidneys, and liver (30). Despite these features, the involvement of the cytotoxin in STEC-associated HUS in humans has not been demonstrated, and the information on the prevalence of SubAB-coding genes in STEC strains associated with severe human disease and in other pathogroups of *E. coli* is still scanty.

During routine Vero cell cytotoxicity assays conducted in our laboratory on *E. coli* strains of human, animal, and food origins, we observed that culture supernatants of two strains isolated from children with uncomplicated diarrhea induced a CPE resembling but distinct from that produced by Stx. The CPE was not neutralized by Stx-neutralizing antisera, and the strains were negative in PCR assays aiming at the identification of Stx-coding genes. In this paper, we demonstrate that the CPE induced by these two strains is governed by SubAB-coding genes and describe for the first time the presence of these genes and the production of the cytotoxin in Stx-negative *E. coli* strains associated with human disease. We also identify the presence of SubAB-coding genes in some LEE-negative STEC strains in our collection and show that, in both Stx-negative and Stx-positive strains, SubAB-coding genes can be located close to the *tia* gene (GenBank accession no. U20318), an invasion genetic determinant previously described in enterotoxigenic *E. coli* (ETEC) (6).

* Corresponding author. Mailing address: Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome, Italy. Phone: 39-06-49903081. Fax: 36-06-49387077. E-mail: stefano.morabito@iss.it.

▽ Published ahead of print on 25 November 2009.

TABLE 1. PCR primers used in this study for *subA* and *subB* gene detection and for characterization of the *subAB* locus

| Primer name | Target gene | Sequence | GenBank accession no. (nucleotide positions) | Reference |
|---|---|---|---|---|
| RTsubABF | *subA* | 5′-GCAGATAAATACCCTTCACTTG-3′ | AF399919 (13856–13835) | 18 |
| RTsubABR | *subB* | 5′-ATCACCAGTCCACTCAGCC-3′ | AF399919 (13625–13643) | 18 |
| tia_lo | *tia* | 5′-TCCATGCGAAGTTGTTATCA-3′ | U20318 (577–558) | This study |
| tia_sense | *tia* promoter | 5′-TTCTCTTTTTACCCTGCTTTTTGC-3′ | FJ664545 (1699–1675) | This study |
| SubAF | *subA* | 5′-GTACGGACTAACAGGGAACTG-3′ | AF399919 (14944–14964) | 18 |
| SubAR | *subA* | 5′-ATCGTCATATGCACCTCCG-3′ | AF399919 (13701–13719) | 18 |
| SubBF | *subB* | 5′-GTAGATAAAGTGACAGAAGGG-3′ | AF399919 (13815–13835) | 18 |
| SubBR | *subB* | 5′-GCAAAAGCCTTCGTGTAGTC-3′ | AF399919 (13121–13140) | 18 |
| SubB_up | *subB* | 5′-TGTATGTGCAGGCGTTATGG-3′ | FJ664545 (3482–3502) | This study |
| subB_end_R | *subB* | 5′-TTATGAGTTCTTTTTCCTGTCAGG-3′ | AF399919 (13283–13306) | This study |
| subA2 | *subA* | 5′-ATGCAATACAGGTGTTTC-3′ | FJ664545 (3101–3083) | This study |

## MATERIALS AND METHODS

**Bacterial strains.** The *E. coli* strains used in this study were part of the culture collection of the Istituto Superiore di Sanità, and many of them have been described in previous studies (12, 13, 16). Serotyping of SubAB-positive strains was performed by the International *Escherichia* and *Klebsiella* Centre (WHO) at the Statens Seruminstitut (Copenhagen, Denmark) or by the Laboratory of Gastrointestinal Pathogens at the Health Protection Agency (London, United Kingdom). Twenty-six *E. coli* strains from healthy human subjects that are included in the ECOR collection were also included in the study (14).

**Detection of cytotoxin production.** The *E. coli* strains were grown in tryptone soy broth (TSB) (Oxoid, Garbagnate Milanese, Italy) at 37°C for 18 h, and culture supernatants were tested by the Vero cell cytotoxicity assay as previously described (3). The cytotoxic titer of a supernatant was defined as the highest dilution inducing CPE after incubation with the Vero cell monolayer for 3 days at 37°C. The presence of Stx was assessed using Stx-neutralizing antisera prepared in rabbits as previously described (3).

**Detection of virulence-associated genes.** The presence of virulence genes associated with the main pathogroups of diarrheagenic *E. coli* strains was assessed by PCR. The intimin-coding *eae* gene was detected as previously described (16); $stx_1$ and $stx_2$ were detected according to Rüssmann et al. (22) and Persson et al. (21), respectively. The other virulence genes considered were the enterohaemolysin-coding gene, *e-hly* (24); *saa* (17), the catalase-peroxidase *katP* (1); the type II secretion system-associated gene, *etpD* (25); the invasion plasmid antigen-coding gene, *ipaH* (11); the enteroaggregative *E. coli* antiaggregation protein transporter gene, *aat* (previously reported as CVD 432) (26); heat-stable enterotoxin 1-coding gene, *astA* (23); the STEC O157 gene encoding the lymphostatin homologue, *toxB* (29); and the enteropathogenic *E. coli* factor for adherence, *efa1* (13). Due to the large dimensions of *toxB* and *efa1*, the 5′ and the 3′ regions of both genes were separately amplified. The heat-labile and heat-stable enterotoxin-coding genes (*elt* and *est*, respectively) were amplified using the primer pairs LTH1/2 and STI1/2, respectively (5, 15). The primer pairs used and the amplification conditions were those described in the respective papers. Amplification products were analyzed on ethidium bromide-stained agarose gels. Primer pairs used for detection of *subA*, *subB*, and *tia* and for the characterization of the *subAB* locus are listed in Table 1.

Long PCR experiments were carried out using the TripleMaster PCR system (Eppendorf, Hamburg, Germany) under the conditions described by the supplier.

**Construction of the EZ::TN<Kan2> mutants.** A library of mutants was prepared from strain ED 591 by random transposon mutagenesis using the EZ::TN <Kan2>Tnp transposome kit (Epicentre, Madison, Wisconsin) under the conditions described in the manual supplied.

**Screening of the EZ::TN<Kan2> mutants library for toxin production.** Single mutant colonies were inoculated into the wells of microtiter plates containing 150 μl of TSB (Oxoid, Garbagnate Milanese, Italy) supplemented with 50 μg/ml kanamycin and incubated at 37°C for 18 h. Plates were then centrifuged at 2,250 × *g* in an Eppendorf 5804R centrifuge, and 20 μl of supernatant from each well were inoculated onto Vero cell monolayers as described elsewhere (3).

**Identification of the genes governing the induction of the CPE.** The mutant negative for toxin production in the Vero cell assay was subjected to genomic sequencing to identify the insertion site of the EZ::TN<Kan2> transposon. Total DNA was prepared from a 5-ml liquid culture by using the Puregene genomic DNA purification kit (Gentra Systems, Big Lake, MN) under the conditions indicated in the manual supplied. Sequencing was performed using the ABI Prism Big Dye terminator cycle sequencing kit version 1.1 (Applied Biosystems, Foster city, CA), according to the manufacturer's instructions. The primers used for sequencing (10) were supplied in the EZ::TN<Kan2>Tnp transposome kit (KAN-2 FP-1 and KAN-2 RP-1). Analyses of sequence data were made using the DNASIS MAX software version 2.0 (Hitachi Software Engineering Co., Ltd.).

**Construction and screening of the genomic library.** A genomic library of strain ED 591 was constructed by using a Lambda ZAP II predigested EcoRI/CIAP-treated vector kit (Stratagene, La Jolla, CA), under the conditions described by the supplier. In detail, 200 ng of total DNA were extracted from strain ED 591 (PureGene; Gentra Systems, Big Lake, MN), partially digested with the EcoRI restriction enzyme, and ligated with 1 μg of the vector by incubation with 2 units of DNA ligase for 18 h at 4°C. Two microliters of the Gigapack III gold packaging extract (Stratagene, La Jolla, CA) were added to the ligase reaction mixture and incubated at room temperature for 2 h. Following addition of 500 μl of SM buffer (5.8 g/liter NaCl, 2.0 g/liter $MgSO_4 \cdot 7H_2O$, 50 mM Tris-HCl [pH 7.5], 0.01% [wt/vol] gelatin) and 20 μl of chloroform, the supernatant containing the phages was titrated in *E. coli* XL1-Blue MRF′ host bacteria. Amplification of the library was performed to obtain a final titer of $1.5 \times 10^9$ PFU/ml.

Plaques were screened using the PCR product obtained with the RTSubABF/RTSubABR primers (Table 1) as a probe. Sequencing of the positive clone was performed using the ABI Prism Big Dye terminator cycle sequencing kit version 1.1 (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions.

**Nucleotide sequence accession number.** A DNA sequence of 3,888 bp comprising the complete sequences of SubA- and SubB-coding genes and part of the *tia* locus from *Escherichia coli* strain ED 591 has been deposited into GenBank under the accession number FJ664545.

## RESULTS

Between 1990 and 2008, more than 3,000 *E. coli* strains of human, animal, and food origins were tested in our laboratory for Stx production by using the Vero cell cytotoxicity assay. The culture supernatants of two strains, termed ED 32 and ED 591, induced the death of the cell cultures, but the CPE was not neutralized by antisera against Stx1 and Stx2. The strains had been isolated from children less than 3 years of age with uncomplicated diarrhea, observed in two different Italian regions in 1990 and 2005, respectively. In both cases, stool samples had been examined for common enteric pathogens (*Salmonella*, *Shigella*, *Campylobacter*, rotavirus) with negative results. The isolates were examined for the presence of Stx-coding genes by PCR, using a panel of primer pairs able to amplify $stx_1$ and the variants of $stx_2$ (21, 22). Both strains gave negative results, thus confirming that the cytotoxin responsible for the CPE did not belong to the Stx family.

**Characterization of the CPE.** Microscopically, the appearance of the CPE was slightly different from that induced by Stx

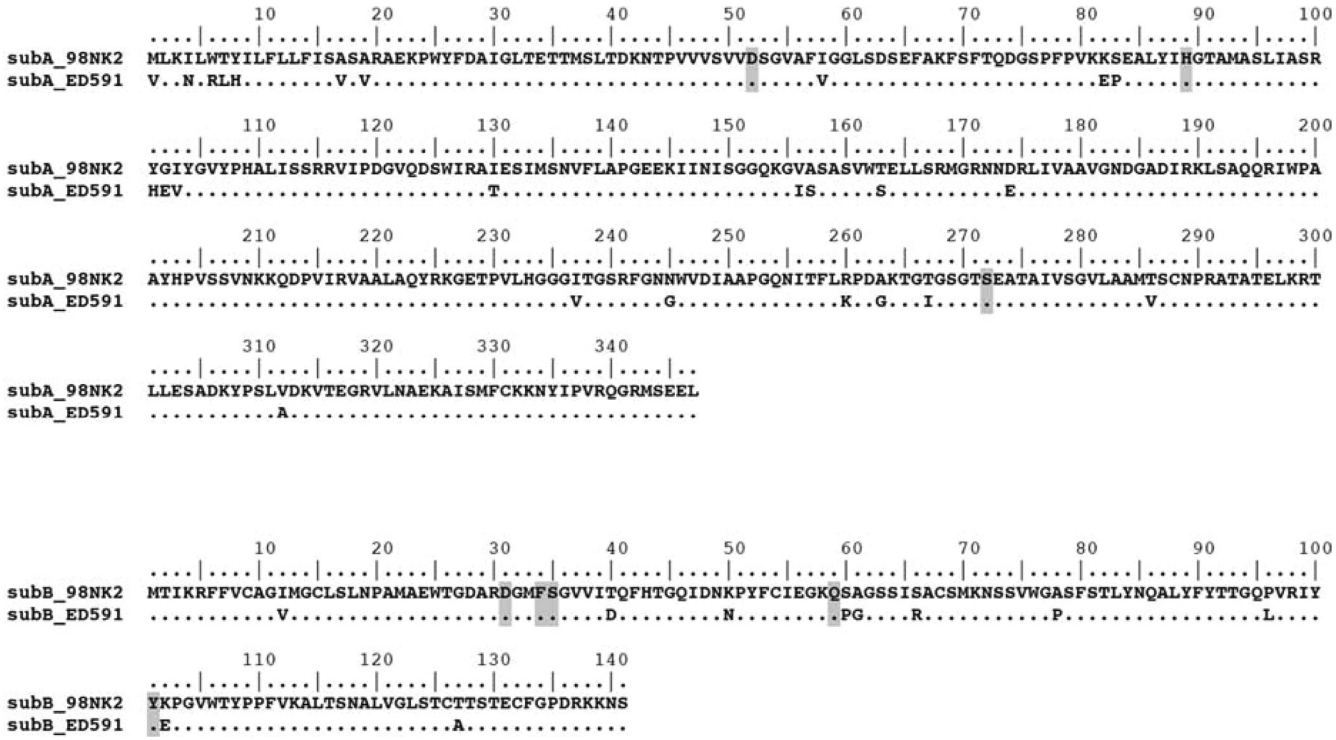



FIG. 1. ClustalW comparison of the amino acidic sequences of the A and B subunits of SubAB from strain ED 591 (this study) and the reference strain 98NK2. The amino acidic substitutions are indicated by the corresponding letters. The residues involved in the catalytic activity in the A subunit and those forming the receptor-binding pocket in the B subunit are indicated (shaded gray).

(not shown). The CPE appeared after 48 h of incubation and advanced by time, whereas the CPE induced by Stx was already visible after 24 h. Substitution of the cell culture medium 6 h after inoculation did not affect the outcome of the CPE. The culture supernatants of both strains induced CPE up to a 1:128 dilution. The toxic activity in culture filtrates of the positive strains was destroyed by heating at 75°C for 15 min or at 60°C for 30 min.

**Identification of the genetic determinant responsible for CPE.** The CPE observed on Vero cells and the absence of Stx-coding genes suggested the possibility that ED 32 and ED 591 strains were producing the recently described SubAB cytotoxin (19). Therefore, they were tested for the presence of SubAB-coding genes by PCR, using the primer pairs (Table 1) described by Paton and colleagues (19). Both strains gave positive results with the primer pair RTsubABF/RTsubABR (Table 1), which is able to amplify a region bridging the end of *subA* and the beginning of *subB*. However, they were negative in PCRs performed with the primer pairs SubAF/SubAR and SubBF/SubBR (Table 1), deployed to amplify the whole *subA* and *subB* genes, respectively.

To demonstrate that the CPE observed was due to a SubAB-like cytotoxin, a library of mutants was prepared from the ED 591 strain by random mutagenesis using the EZ::TN<Kan2> transposon. One thousand eight hundred mutants able to grow on media containing 50 μg/ml kanamycin were obtained and tested by the Vero cell cytotoxicity assay. One of the mutants proved negative and was further analyzed to determine the insertion site of the transposon. Sequence analysis showed that the transposon was located within the coding sequence of the *subA* gene, thus demonstrating the connection between the CPE observed on Vero cells and the presence of a functional *subA* gene.

**Identification and characterization of the DNA region containing the SubAB-coding genes.** A genomic library obtained from total DNA of the ED 591 strain was screened by Southern hybridization using the RTsubABF/RTsubABR PCR product as a probe. One clone of about 3.7 kb in length was positive, and sequence analysis showed that it contained the complete *subA* gene and most of the *subB* gene at one extremity. The sequence of the whole *subB* gene was obtained by PCR amplification of total DNA from the ED 591 strain, using the primer pair SubB_up/SubB_end_R deployed on the *subB* gene sequence from pO113 (GenBank accession no. AF399919) (Table 1). The alignment of the *subA* and *subB* sequences from ED 591 with the corresponding sequences in pO113 returned a homology of 90% at the nucleotide level. The alignment of the translated amino acid sequences showed 93% identity and 96% positive substitutions for the SubA protein, and 93% identity and 94% positive substitutions for the SubB protein (Fig. 1). The 3.7-kb clone also contained a portion of the *tia* gene, encoding an invasion determinant described in ETEC strains (6), interrupted by an IS*2* element (Fig. 2). Consistently, a *tia*-specific PCR amplification conducted using total DNA from the ED 591 strain with the primer pair tia_sense/tia_lo (Table 1 and Fig. 1) yielded an amplicon of about 1.8 kb (Fig. 3), the size expected for the 650-bp *tia* sequence (GenBank accession no. U20318) interrupted by a 1.2-kb IS*2* element. This DNA trait, located about 800 bp upstream of *subA*, is not present in the sequence of the pO113 (GenBank

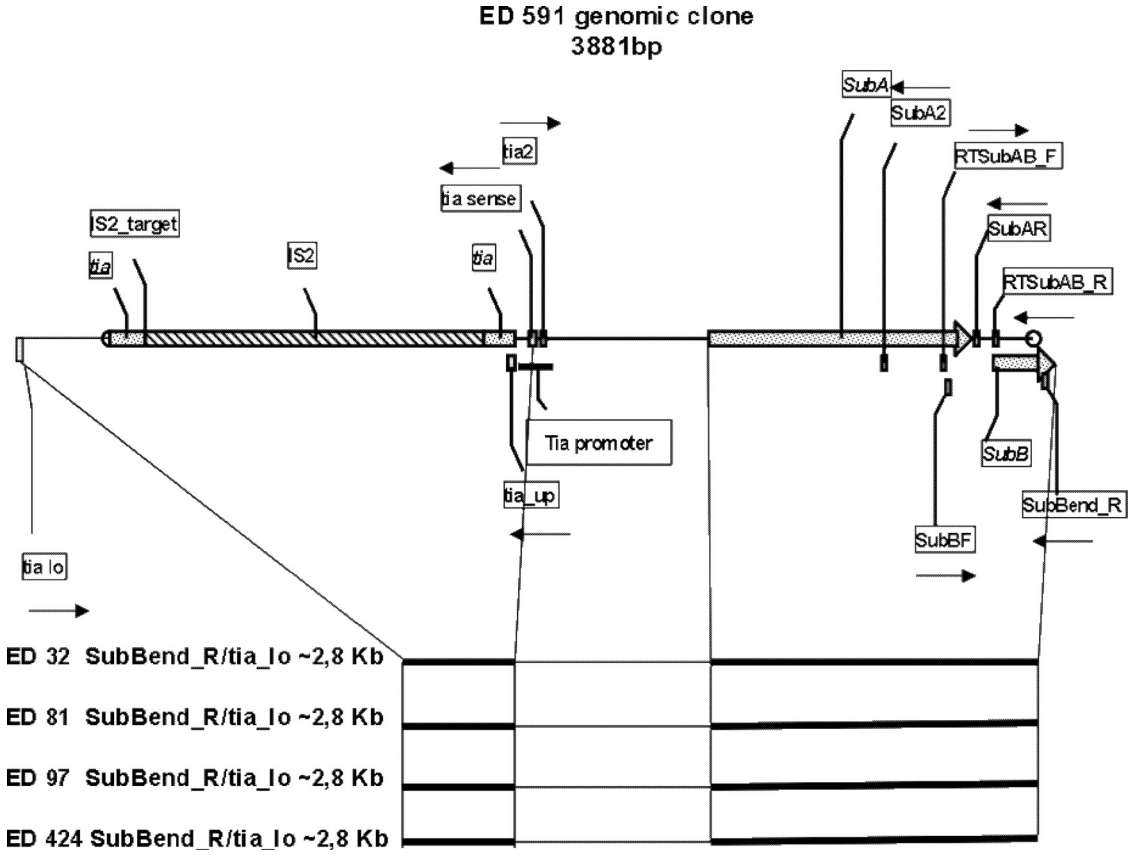



FIG. 2. Molecular architecture of the *subAB* locus in the *subAB-tia*-positive strains. The primers' positions and orientations are indicated by arrows.

accession no. AF399919), suggesting that, in ED 591, *subAB* may be located in a different genomic region.

**Characterization of the *subAB* locus in strain ED 32.** A set of PCR primers was used to characterize the locus encoding the SubAB toxin of strain ED 32 (Fig. 2 and Table 1). The primers were either described in the literature or deployed on the sequence of the *subAB* locus obtained for strain ED 591. PCR mapping showed that the locus encoding the SubAB cytotoxin of strain ED 32 was similar to the locus present in strain ED 591, with the exception that, in the former, the *tia* gene was not interrupted by the IS*2* element (Fig. 2). This finding was confirmed by sequencing the PCR product obtained from strain ED 32 by using the primers SubA2 and tia_lo (Fig. 2).

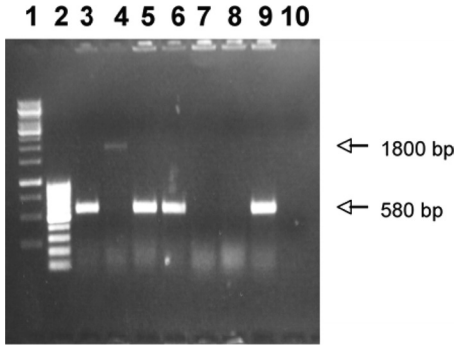



FIG. 3. PCR amplification of *tia* in strains ED 32, ED 591, ED 81, ED 97, ED 99, ED 186, and ED 424 and in the negative control (lanes 3 to 10, respectively) using the primer pair tia_sense/tia_lo. Molecular weight markers are indicated in lanes 1 and 2 (1 kb and 100 bp, respectively). Refer to Fig. 1 and Table 1 for primer location.

**Identification of other *subAB*-positive strains.** A panel of 61 Stx-positive and 74 Stx-negative *E. coli* strains from the culture collection of our laboratory was examined for the presence of *subAB* genes by PCR, using the primer pair RTsubABF/RTsubABR. The STEC strains belonged to serogroups O157 (17 strains), O26 (12 strains), O111 (eight strains), O103 (three strains), and O113 (eight strains) and to 10 other different serogroups (13 strains). The Stx-negative strains included 43 enteropathogenic *E. coli*, 10 ETEC, 13 enteroaggregative *E. coli*, and eight cytotoxic necrotizing factor 1-producing *E. coli* strains. Twenty-six nonpathogenic *E. coli* strains isolated from healthy humans and that are part of the ECOR collection (14) were also included in the study. All the Stx-negative strains were also negative in the PCR assay, while five STEC strains of human origin, namely ED 81, ED 97, ED 99, ED 186, and ED 424 (Table 2), gave positive results. Strains ED 81 and ED 424 had been isolated from cases of uncomplicated diarrhea, while the remaining were from an HUS case (ED 186) and from household contacts of cases of HUS (ED 97 and ED 99).

TABLE 2. Serotypes and virulence genes of the *subAB*-positive *E. coli* strains

| Strain (serotype) | Presence or absence of[a]: | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *eae* | *stx*$_1$ | *stx*$_2$ | *tia* | *aat* | *e-hly* | *katP* | *toxB* | *efa1* | *saa* | *etpD* | *astA* | *ial* | *elt* | *est* |
| ED 32 (O78:H−) | − | − | − | + | − | + | − | − | − | − | − | − | − | − | − |
| ED 591 (O8:Hnd[b]) | − | − | − | + | − | − | − | − | − | − | − | − | − | − | − |
| ED 81 (O128:H2) | − | − | + | + | − | + | − | − | − | − | − | − | − | − | − |
| ED 424 (O113:Hnd[b]) | − | + | + | + | − | + | − | − | − | − | − | − | − | − | − |
| ED 99 (O120:H19) | − | + | + | − | − | + | − | − | − | + | − | − | − | − | − |
| ED 186 (O48:H21) | − | + | − | − | − | − | + | − | − | + | − | + | − | − | − |
| ED 97 (O91:H−) | − | + | − | + | − | − | − | − | − | − | − | − | − | − | − |

[a] +, present; −, absent.
[b] nd, not determined.

**Characterization of *subAB*-positive strains.** The seven *subAB*-positive strains were investigated by PCR for the presence of *tia* and other virulence genes associated with diarrheagenic *E. coli*, and their features are reported in Table 2. All the strains were negative for the intimin-coding *eae* gene, but four possessed the enterohemolysin-coding gene, which is considered a hallmark of the presence of large virulence plasmids associated with STEC O157 and other pathogenic STEC strains (9). However, these isolates were negative in PCR assays specific for other virulence determinants often located on those plasmids (1, 25). Besides the two Stx-negative strains, three of the five STEC strains were positive for *tia*, which was also present in only two of the 56 *subAB*-negative STEC strains, belonging to serogroups O111 and O103. All the *subAB-tia*-positive strains were also positive in long PCR experiments conducted using the primer pair subB_end_R/tia_lo, yielding amplicons of about 2.8 kb. This result indicated that *subAB* and *tia* were located contiguously in all the strains with an architecture similar to that observed in strain ED 591, apart from the presence of the IS*2* element interrupting *tia* in the latter (Fig. 2). The two *tia*-negative strains (ED 99 and ED 186) were positive for the *saa* locus, which was conversely absent in the *tia*-positive strains.

## DISCUSSION

SubAB produced by *E. coli* is an $AB_5$ toxin able to induce a CPE on Vero cells (19) and the typical features of the Stx-induced HUS in mice (30). Till now, the production of SubAB or the presence of *subAB* genes has been reported only for certain STEC strains (4, 18, 19). This paper, to the best of our knowledge, represents the first description of the production of SubAB and of the presence of *subAB* genes in Stx-negative strains of *E. coli*. Our data also confirm that SubAB cytotoxins are active on Vero cells (19), since we showed that the CPE induced by culture supernatants of the Stx-negative strain ED 591 was suppressed by the interruption of the *subA* gene by insertional mutagenesis.

Genetic analyses of the SubAB-positive, Stx-negative strains, ED 32 and ED 591, showed that the nucleotide sequences of the *subA* and *subB* genes were identical in the two strains and 90% similar to that of the corresponding genes present in the pO113 plasmid of strain 98NK2 (19). We do not know if this sequence polymorphism may reflect differences in the biological properties of the toxin. However, it does not have implications for diagnostics, since the PCR primers designed by Paton and colleagues (18, 19) on the sequence obtained from the prototype strain 98NK2 were able to identify the presence of the *subAB* genes of the Stx-negative strains described in this study.

All the SubAB-positive STEC strains described so far, despite belonging to different serogroups, had in common the lack of the LEE locus (4, 18) and, in many cases, the presence of *saa*, a locus encoding the production of an autoagglutinating adhesin possibly involved in the colonization of the host intestinal mucosa (20). Both strains ED 32 and ED 591 were negative for *saa*, while their *subAB* genes were contiguous to a gene homologous to *tia*, a genetic determinant responsible for the production of a 25-kDa outer membrane protein required for the invasion of intestinal epithelial cells, described in ETEC strains and associated with the pathogenesis of ETEC-induced disease (6). The *tia* gene is not present in the sequence of the pO113 plasmid (GenBank accession no. AF399919), and this suggests that the *subAB* genes of the Stx-negative strains described in this study may be framed by a different DNA region.

PCR screening of a collection of *E. coli* strains belonging to different pathogroups and of nonpathogenic strains from the ECOR collection (14) revealed the presence of *subAB* only in five STEC strains. Since the strains tested were selected to represent different serogroups and virulence gene assets, no conclusions can be drawn on the prevalence of SubAB in STEC isolated in Italy. On the other hand, the negative PCR results, together with those from the routine use of the Vero cell assays performed in the past years in our laboratory, indicate that SubAB production is rare among Stx-negative strains, even if further studies are needed to clarify this issue.

As far as the presence of additional virulence genes was concerned, all the *subAB*-positive strains detected were negative in PCR assays specific for the intimin-coding *eae* gene, thus confirming that the presence of *subAB* appears to be restricted to LEE-negative *E. coli* strains (4, 18). However, in a difference from previous reports (4, 18, 19), we found the presence of the *saa* gene in only two of the five SubAB-positive STEC strains identified in this study, while the remaining three were positive for the *tia* gene, as were the two Stx-negative strains. The presence of *tia* has been occasionally reported in EPEC and enteroaggregative *E. coli* (6); however, it does not appear to be a common feature in STEC, since it has been detected in only two of the 56 *subAB*-negative STEC strains investigated in this study. PCR mapping and sequence analyses also showed that

in the three STEC strains, *tia* and *subAB* were physically associated in the same DNA region (Fig. 2), which could represent a putative pathogenicity island (PAI) (8). This assumption is supported by the GC content of the DNA region (48.4%) (GenBank accession no. FJ664545), which is lower than the average GC content estimated for the *E. coli* K-12 genome (50.8%).

All the *tia*-positive strains did not possess *saa*, which was conversely present in the two *tia*-negative STEC strains (Table 2). Since *saa* has been described in the pO113 plasmid, this observation, together with the absence of *tia* in the pO113 nucleotide sequence, suggests that the putative *subAB-tia* PAI is not located in this type of plasmid. This assumption has been confirmed by Southern hybridization experiments showing that a *subAB* probe, prepared by PCR using the RTsubABF/RTsubABR primer pair, did not hybridize with the purified plasmids obtained from strains ED32 and ED 591 but gave positive signals with their total DNA (data not shown).

In conclusion, we provide the first evidence that SubAB cytotoxins can be produced by Stx-negative *E. coli* and can be associated with human disease even in the absence of Stx production by the infecting strain. The nucleotide sequences of the cytotoxins produced by the two strains identified in this study were identical but slightly different from those of the prototype SubAB described by Paton et al. (19). Moreover, the *subAB* genes of the Stx-negative strains were associated with *tia* and not with *saa*, which was present in many of the *subAB*-positive STEC strains described so far (2, 4). The *subAB* genes were contiguous to *tia* also in three of the five *subAB*-positive STEC strains identified in this study. The observation of this new gene cluster suggests that, alternative to plasmid pO113, *subAB* genes may be vehiculated by a putative PAI and that multiple evolutionary events may have occurred during the dissemination of these genes.

## ACKNOWLEDGMENTS

This work was partially supported by grants from the European Community Network of Excellence Med-Vet-Net (contract FOOD-CT-2004-506122) and from the Italian Ministry of University and Research (SAFE-EAT project).

## REFERENCES

1. **Brunder, W., H. Schmidt, and H. Karch.** 1996. KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic *Escherichia coli* O157:H7. Microbiology **142:**3305–3315.
2. **Byres, E., A. W. Paton, J. C. Paton, J. C. Löfling, D. F. Smith, M. C. J. Wilce, U. M. Talbot, D. C. Chong, H. Yu, S. Huang, X. Chen, N. M. Varki, A. Varki, J. Rossjohn, and T. Beddoe.** 2008. Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin. Nature **456:**648–652.
3. **Caprioli, A., I. Luzzi, F. Rosmini, P. Pasquini, R. Cirrincione, A. Gianviti, M. C. Matteucci, and G. Rizzoni.** 1992. Hemolytic-uremic syndrome and Vero cytotoxin-producing *Escherichia coli* infection in Italy. J. Infect. Dis. **166:**154–158.
4. **Cergole-Novella, M. C., L. S. Nishimura, L. F. Dos Santos, K. Irino, T. M. Vaz, A. M. Bergamini, and B. E. Guth.** 2007. Distribution of virulence profiles related to new toxins and putative adhesins in Shiga toxin-producing *Escherichia coli* isolated from diverse sources in Brazil. FEMS Microbiol. Lett. **274:**329–334.
5. **Dallas, W.** 1990. The heat-stable toxin I gene from *Escherichia coli* 18D. J. Bacteriol. **172:**5490–5493.
6. **Fleckenstein, J. M., D. J. Kopecko, R. L. Warren, and E. A. Elsinghorst.** 1996. Molecular characterization of the *tia* invasion locus from enterotoxigenic *Escherichia coli*. Infect. Immun. **64:**2256–2265.
7. **Frankel, G., A. D. Phillips, I. Rosenshine, G. Dougan, J. B. Kaper, and S. Knutton.** 1998. Enteropathogenic and enterohaemorrhagic *Escherichia coli*: more subversive elements. Mol. Microbiol. **30:**911–921.
8. **Hochhut, B., U. Dobrindt, and J. Hacker.** 2005. Pathogenicity islands and their role in bacterial virulence and survival. Contrib. Microbiol. **12:**234–254.
9. **Karch, H., H. Schmidt, and W. Brunder.** 1998. Plasmid encoded determinants of *Escherichia coli* O157:H7, p. 183–194. *In* J. Kaper and A. D. O'Brien (ed.), *Escherichia coli* O157:H7 and other Shiga toxin-producing *E. coli* strains. American Society for Microbiology, Washington, DC.
10. **Kirby, J. R.** 2007. *In vivo* mutagenesis using EZ-Tn*5*. Methods Enzymol. **421:**17–21.
11. **Lüscher, D., and M. Altwegg.** 1994. Detection of shigellae, enteroinvasive and enterotoxigenic *Escherichia coli* using the polymerase chain reaction (PCR) in patients returning from tropical countries. Mol. Cell Probes **8:**285–290.
12. **Morabito, S., R. Tozzoli, A. Caprioli, H. Karch, and A. Carattoli.** 2002. Detection and characterization of class 1 integrons in enterohemorrhagic *Escherichia coli*. Microb. Drug Resist. **8:**85–91.
13. **Morabito, S., R. Tozzoli, E. Oswald, and A. Caprioli.** 2003. A mosaic pathogenicity island made up of the locus of enterocyte effacement and a pathogenicity island of *Escherichia coli* O157:H7 is frequently present in attaching and effacing *E. coli*. Infect. Immun. **71:**3343–3348.
14. **Ochman, H., and R. K. Selander.** 1984. Standard reference strains of *Escherichia coli* from natural populations. J. Bacteriol. **157:**690–693.
15. **Olsvik, O., and N. A. Strockbine.** 1993. PCR detection of heat-stable, heat-labile, and Shiga-like toxin genes in *Escherichia coli*, p. 271–276. *In* D. H. Persing, T. F. Smith, F. C. Tenover, and T. J. White (ed.), Diagnostic molecular microbiology. Principles and applications. American Society for Microbiology, Washington, DC.
16. **Oswald, E., H. Schmidt, S. Morabito, H. Karch, O. Marches, and A. Caprioli.** 2000. Typing of intimin genes in human and animal enterohemorrhagic and enteropathogenic *Escherichia coli*: characterization of a new intimin variant. Infect. Immun. **68:**64–71.
17. **Paton, A. W., and J. C. Paton.** 2002. Direct detection and characterization of Shiga toxigenic *Escherichia coli* by multiplex PCR for *stx1*, *stx2*, *eae*, *ehxA*, and *saa*. J. Clin. Microbiol. **40:**271–274.
18. **Paton, A. W., and J. C. Paton.** 2005. Multiplex PCR for direct detection of Shiga toxigenic *Escherichia coli* strains producing the novel subtilase cytotoxin. J. Clin. Microbiol. **43:**2944–2947.
19. **Paton, A. W., P. Srimanote, U. M. Talbot, H. Wang, and J. C. Paton.** 2004. A new family of potent AB(5) cytotoxins produced by Shiga toxigenic *Escherichia coli*. J. Exp. Med. **200:**35–46.
20. **Paton, A. W., P. Srimanote, M. C. Woodrow, and J. C. Paton.** 2001. Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. Infect. Immun. **69:**6999–7009.
21. **Persson, S., K. E. Olsen, S. Ethelberg, and F. Scheutz.** 2007. Subtyping method for *Escherichia coli* Shiga toxin (verocytotoxin) 2 variants and correlations to clinical manifestations. J. Clin. Microbiol. **45:**2020–2024.
22. **Rüssmann, H., E. Kothe, H. Schmidt, S. Franke, D. Harmsen, A. Caprioli, and H. Karch.** 1995. Genotyping of Shiga-like toxin genes in non-O157 *Escherichia coli* strains associated with haemolytic uraemic syndrome. J. Med. Microbiol. **42:**404–410.
23. **Savarino, S. J., A. Fasano, J. Watson, B. M. Martin, M. M. Levine, S. Guandalini, and P. Guerry.** 1993. Enteroaggregative *Escherichia coli* heat-stable enterotoxin 1 represents another subfamily of *E. coli* heat-stable toxin. Proc. Natl. Acad. Sci. USA **90:**3093–3097.
24. **Schmidt, H., L. Beutin, and H. Karch.** 1995. Molecular analysis of the plasmid-encoded hemolysin of *Escherichia coli* O157:H7 strain EDL933. Infect. Immun. **63:**1055–1061.
25. **Schmidt, H., B. Henkel, and H. Karch.** 1997. A gene cluster closely related to type II secretion pathway operons of Gram-negative bacteria is located in the large plasmid of enterohemorrhagic *Escherichia coli* O157 strains. FEMS Microbiol. Lett. **148:**265–272.
26. **Schmidt, H., C. Knop, S. Franke, S. Aleksic, J. Heesemann, and H. Karch.** 1995. Development of PCR for screening of enteroaggregative *Escherichia coli*. J. Clin. Microbiol. **33:**701–705.
27. **Spangler, B. D.** 1992. Structure and function of cholera toxin and the related *Escherichia coli* heat-labile enterotoxin. Microbiol. Rev. **56:**622–647.
28. **Tarr, P. I., C. A. Gordon, and W. L. Chandler.** 2005. Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. Lancet **365:**1073–1086.
29. **Tozzoli, R., A. Caprioli, and S. Morabito.** 2005. Detection of *toxB*, a plasmid virulence gene of *Escherichia coli* O157, in enterohemorrhagic and enteropathogenic *E. coli*. J. Clin. Microbiol. **43:**4052–4056.
30. **Wang, H., J. C. Paton, and A. W. Paton.** 2007. Pathologic changes in mice induced by subtilase cytotoxin, a potent new *Escherichia coli* $AB_5$ toxin that targets the endoplasmic reticulum. J. Infect. Dis. **196:**1093–1101.

# A new pathogenicity island carrying an allelic variant of the Subtilase cytotoxin is common among Shiga toxin producing *Escherichia coli* of human and ovine origin

Running title: *subAB* allelic variants in STEC strains

Keywords: STEC, Pathogenicity island, Subtilase cytotoxin, diarrhoea, small ruminants, PCR

*Valeria Michelacci[1,2], Rosangela Tozzoli[1], Alfredo Caprioli[1], Remigio Martínez[3], Flemming Scheutz[4], Laura Grande[1], Sergio Sánchez[5] and Stefano Morabito[1#]*

[1] European Reference Laboratory for *Escherichia coli*, Istituto Superiore di Sanità, Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare, Rome, Italy

[2] Department of Biology, University "Roma Tre", Roma, Italy

[3] Patologia Infecciosa y Epidemiologia, Departamento de Sanidad Animal, Facultad de Veterinaria, Universidad de Extremadura, Cáceres, Spain

[4] WHO Collaborating Centre for Reference and Research on *Escherichia* and *Klebsiella*, Department of Microbiology & Infection Control, Statens Serum Institut, Copenhagen, Denmark

[5] Laboratorio de Enterobacterias, Servicio de Bacteriologia, Centro Nacional de Microbiologia, Instituto de Salud Carlos III, Madrid, Spain

[#]Corresponding author. Mailing address:
Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare,
Istituto Superiore di Sanità,
Viale Regina Elena 299, 00161, Rome, Italy
Tel. +39-06-49903081
Fax +36-06-49387077
E-mail stefano.morabito@iss.it

27

38

## Abstract

Subtilase (SubAB) is a cytotoxin elaborated by some Shiga Toxin (Stx)-producing *Escherichia coli* (STEC) strains usually lacking the locus of enterocyte effacement (LEE). Two variants of SubAB coding genes have been described: *subAB*$_1$, located on the plasmid of the STEC O113 98NK2 strain and *subAB*$_2$, located on a pathogenicity island (PAI) together with the *tia* gene, encoding an invasion determinant described in enterotoxigenic *E. coli*.

In the present study, we determined the entire nucleotide sequence of the PAI containing the *subAB*$_2$ operon, termed Subtilase-Encoding PAI (SE-PAI), and identified its integration site in the *pheV* tRNA locus. In addition, a PCR strategy for discriminating the two *subAB* allelic variants was developed and used to investigate their presence in *E. coli* strains belonging to different pathotypes and in a large collection of LEE-negative STEC of human and ovine origin. The results confirmed that *subAB* genes are carried predominantly by STEC and showed their presence in 72% and 86% of the LEE-negative strains from human cases of diarrhoea and from healthy sheep examined, respectively. Most of the *subAB*-positive strains (98%) identified possessed the *subAB*$_2$ allelic variant and were also positive for *tia,* suggesting the presence of SE-PAI. Altogether, our observations indicate that *subAB*$_2$ is the prevalent SubAB-coding operon in LEE-negative STEC circulating in European countries, and that sheep may represent an important reservoir for human infections with these strains. Further studies are needed to assess the role of *tia* and/or other genes carried by SE-PAI in the colonization of the host intestinal mucosa.

## Introduction

Subtilase (SubAB) is an $AB_5$ toxin produced by certain *Escherichia coli* strains associated with human disease [1]. SubAB is constituted by a 35 kDa A subunit displaying a subtilase-like serine protease activity and by five 13 kDa B subunits forming a pentamer, which mediates the binding to specific receptors on the host cell surface [1]. Following internalization in cultured cells, SubAB is delivered to the endoplasmic reticulum (ER) [2] where it has been demonstrated to cleave the chaperone BiP [3], causing the RNA-dependent protein kinase-like ER kinase activation and the transient inhibition of protein synthesis, resulting in the induction of the apoptotic signalling pathways [4, 5, 6, 7].

SubAB has been so far identified almost exclusively in Shiga toxin (Stx)-producing *E. coli* (STEC), and in particular in strains that do not possess the locus for enterocyte effacement (LEE) [8-18]. The LEE is a pathogenicity island (PAI) governing the attaching and effacing mechanism of intestinal adhesion [19], and represents a common feature of STEC strains associated with severe human disease. It has been hypothesized that the SubAB may contribute to the pathogenesis of STEC-associated human disease by playing a synergistic action with Stx [1]. As a matter of fact, SubAB has been shown to induce, in a mouse model, the typical haemolytic uremic syndrome (HUS)-associated features caused by Stx, such as extensive microvascular damage, and thrombosis and necrosis in brain, kidneys, and liver [20].

The prototype SubAB-positive STEC strain 98NK2, belonging to serotype O113:H21 and isolated from an outbreak of HUS in South Australia [1], carries the subtilase-coding operon (*subAB*) on a large virulence plasmid designated as pO113, which also carries the *saa* gene, encoding an autoagglutinating adhesin possibly involved in the colonization of the host intestinal mucosa [21].

Recently, we reported the production of SubAB by two Stx-negative *E. coli* strains (ED 32 and ED 591), isolated from two unrelated cases of uncomplicated diarrhoea in Italy [22].

Genetic analyses showed that the nucleotidic sequences of the *subA* and *subB* genes were identical in the two strains and 90% similar to those of the corresponding genes present in the pO113 plasmid of strain 98NK2 [22]. Strains ED 32 and ED 591 were both LEE-negative and did not react in a *saa*-specific PCR assay. Moreover, differently from strain 98NK2, they harboured the *subAB* genes in the chromosome and next to another gene, *tia*, encoding an invasion factor previously described in enterotoxigenic *E. coli* (ETEC) [23]. An identical chromosomal region carrying *subAB* and *tia* was identified in the chromosome of other *subAB*-positive STEC strains [22], suggesting the existence of a putative pathogenicity island (PAI) vehiculating the *subAB* and *tia* virulence genes.

The presence of *subAB* genes among *E. coli* strains of human or animal origin has been investigated in several studies [8-18]. However, most of these studies did not involve the use of tools capable to distinguish between the two allelic variants of the *subAB* gene. As an exception, a recent investigation conducted on STEC strains isolated from cattle, sheep and goats [14] reported a different distribution of the two *subAB* variants in the different animal species, with the $subAB_{98NK2}$, named by the authors $subAB_1$, associated with bovine strains, and $subAB_{ED32}$, termed $subAB_2$, more frequent among strains from small ruminants [14].

In the present work, we investigated the presence of the two allelic variants of the *SubAB* gene in human *E. coli* strains belonging to different pathotypes and in a large collection of LEE-negative STEC of human and ovine origin. Moreover, we determined the entire nucleotidic sequence of the putative PAI containing the $subAB_2$ operon in the prototype *E. coli* strain ED 32 and describe its gene content and insertion site.

## Materials and Methods

### *Bacterial strains*

The prototype *E. coli* strain ED 32, containing the *subAB*$_2$ allelic variant, is part of the culture collection of the Istituto Superiore di Sanità and has been previously described [22]. The panel of human strains investigated included 177 STEC strains belonging to 10 different serogroups and displaying different combinations of *stx*-coding genes, 26 enteropathogenic *E. coli* (EPEC), 32 enterotoxigenic *E. coli* (ETEC), 20 enteroaggregative *E. coli* (EAEC), 13 enteroinvasive *E. coli* (EIEC), all isolated from cases of diarrhoea, and one strain isolated from a urinary tract infection (uropathogenic *E. coli*, UPEC).

The diarrheagenic *E. coli* used in this study have been classified based on the presence of the virulence genes described to be associated to the different pathotypes in the literature. EPEC pathotype was identified by the presence of the intimin-coding *eae* gene, detected together with the plasmid-associated marker EAF for identifying typical EPEC. STEC were recognized by the presence of the *stx1* and *stx2* genes. The invasion plasmid antigen-coding gene *ipaH* and the enteroaggregative *E. coli* antiaggregation protein transporter gene, *aat* (previously reported as CVD 432) were considered markers for EIEC and EAEC pathotypes respectively. Finally, ETEC strains were identified by the presence of the heat-stable and the heat-labile enterotoxins-coding genes (*est* and *elt*, respectively). All the virulence genes were amplified as previously described [22]. The PCR assay for the gene encoding the EAF determinant has been described in [24]. The only human UPEC was isolated from a patient suffering from Urinary Tract Infection. Serotyping, including both O and H antigen identification, was performed by the WHO Collaborating Centre for Reference and Research on *Escherichia* and *Klebsiella* at the Statens Serum Institut (Copenhagen, DK).

All the STEC strains selected for this study lacked the *eae* gene, considered as a hallmark for the presence of LEE locus. All the *E. coli* strains of human origin included in the study are part of the Statens Serum Institut (Copenhagen, DK) culture collection.

One hundred and twenty-three LEE-negative STEC strains isolated from sheep were included in the study. The strains had been isolated from healthy animals in Spain during a previous longitudinal study involving 12 different sampling visits (one sampling/month) at four different farms. All the animal isolates are part of the culture collection held at the University of Extremadura (Caceres, ES).

***Construction of a genomic library from ED 32 strain***

A genomic library of strain ED 32 was constructed by using a Lambda ZAP II predigested *Eco*RI/CIAP-treated vector kit (Stratagene, La Jolla, CA), in the conditions described by the supplier. In detail, 200 ng of total DNA were extracted from strain ED 32 (PureGene; Gentra Systems, Big Lake, MN), partially digested with the *Eco*RI restriction enzyme, and ligated with 1 μg of the vector by incubation with 2 units of DNA ligase for 18 h at 4°C. Two μl of the Gigapack III gold packaging extract (Stratagene, La Jolla, CA) were added to the ligase reaction mixture and incubated at room temperature for 2 h. Following addition of 500 μl of SM buffer (5.8 g/liter NaCl, 2.0 g/liter $MgSO_4 \cdot 7H_2O$, 50 mM Tris-HCl [pH 7.5], 0.01% [wt/vol] gelatin) and 20 μl of chloroform, the supernatant containing the phages was titrated in *E. coli* XL1-Blue MRF' host bacteria. Amplification of the library was performed to obtain a final titre of 1.5 X $10^9$ PFU/ml.

***Identification of the subAB-flanking regions in the ED 32 strain***

The sequence of the chromosomal regions flanking the *subAB* and *tia* genes in strain ED 32 was determined by using two different approaches. Since the *tia* gene associated with *subAB*$_2$ shared 90% of sequence identity with the homologous gene in the ETEC strain IAI39, the region downstream this gene was obtained by PCR using the primer *tRNA_phe_anchor* designed on the available genomic sequence of this latter strain (GenBank Acc. No. CU928164), in combination with the *subA_anchor* primer, deployed on

the sequence of the ED 32 strain *subA* gene (Figure 1 and Table 1). The amplification reaction was conducted on a total DNA preparation from strain ED 32 and gave an amplification product of about 6,600 bp, which was subjected to sequencing.

The DNA sequence of the region located downstream the *subAB* operon was determined by sequencing the insert of a clone identified by screening the ED 32 library with a DNA probe corresponding to the 3' terminal part of the *subB* gene obtained by PCR amplification using the subBendF/subBendR primer pair (Table 1) [22 and this study].

***Evaluation of the activity of the two variants of SubAB by Vero Cells Assay***

To investigate on the possible differences in the activity of the two SubAB variants their ability of inducing CPE onto Vero cells monolayers has been compared by Vero cell assay (VCA). The culture supernatant from the *E. coli* strains ED 186 and ED 32, producing the $SubAB_1$ and $SubAB_2$ respectively, has been assayed by inoculating serial dilutions, up to 1:1024, onto Vero cells monolayers in microtiter plates. Since the ED 186 strain also produced the Stx1, each dilution of the culture supernatant from this strain, has been neutralized by incubation with an antibody raised against Stx1, at a working titer of 1:200, before inoculation. The VCA has been conducted as previously described [22] and the CPE has been observed at 24 hours, 48 hours and 72 hours after inoculation. The *subAB* genes from the two strains have also been cloned in the pGEM-T-Easy vector and used to transform the K-12 strain JM109. The *subAB* operons have been cloned together with their native promoters to allow the expression of the subtilase. Serial dilutions of overnight cultures supernatants of the K-12 strains containing the cloned *subAB* operons have been inoculated onto Vero cells monolayer after sonication of the culture to facilitate the release of the toxin produced and analysed in parallel with the wild type strains in the same conditions.

***PCR screening for the presence of subAB, saa, and tia genes***

The presence of *subAB* was assessed using the primer pair RTsubABF/RTsubABR [1], which allows the detection of both the allelic variants [22]. The *subAB* operon from the positive strains was typed using the primer pairs SubAF/RTsubABR [this study and 1] and subA_startF/RTsubABR [this study and 1], able to specifically detect the prototype *subAB* (*subAB*$_1$) or the allelic variant present in the ED 32 strain (*subAB*$_2$), respectively. The presence of *saa* and *tia* genes was assessed as previously described [22]. All the primers used are listed in Table 1.

***Nucleotide sequence accession number***

The DNA sequence of the 8,058 bp DNA stretch spanning the complete sequence of the PAI vehiculating *subAB* in the *E. coli* strain ED 32 was submitted into GenBank with the Acc. No. JQ994271. Annotation has been made through Glimmer online tool (http://www.ncbi.nlm.nih.gov/genomes/MICROBES/glimmer_3.cgi) and the predicted open reading frames have been analysed with blastp algorithm using the tool available at NCBI.. (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi?PROGRAM=blastp&BLAST_PROGRAMS=blastp&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome) in order to investigate on the closest informative matches (Table 2).

All the nucleotidic alignments have been done with the blastn algorithm using the tool available at NCBI.

(http://www.ncbi.nlm.nih.gov/blast/Blast.cgi?PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome)

## Results

***Characterization of the genomic locus harboring the subAB$_2$ operon in strain ED 32***

The nucleotidic sequence of the genomic island vehiculating the *subAB*$_2$ operon in strain ED 32, partially described in a previous study, was completed and the structure of the

locus is schematically reported in figure 1. Sequence analysis showed that the locus, that we termed Subtilase-Encoding PAI (SE-PAI), spans the 8 kb region between the *E. coli* gene *yjhS,* located 314 bp downstream the *subAB*$_2$ operon, and the *pheV* tRNA locus, situated 4,725 bp from the end of the *tia* gene (Figure 1). The latter DNA region contains an unnamed gene encoding a sulfatase, the *shiA* gene, whose product is involved in the attenuation of *Shigella flexneri*-induced inflammatory response, and the genetic determinant for an integrase, probably involved in the PAI mobilization machinery (Table 2). Downstream the *subAB*$_2$ operon, SE-PAI ends before the *yjhS* gene, which is part of an operon, named *yjhATS* or *nanCMS*, encoding factors involved in the internalization and catabolism of sialic acids. Such an operon was described in both commensal and pathogenic bacteria and is also present in the *E. coli* K12 reference strain MG1655 (GenBank Acc. No. NC000913). Therefore, it was considered as the downstream boundary of SE-PAI.

The whole sequence of the SE-PAI has been searched for homology with the records of the nucleotidic sequences, including the whole genomes, stored in GenBank, with aim of investigating the variability of the ORFs and intergenic sequences of the PAI. Three contigs deriving from whole genome sequencing projects of STEC (*E. coli* 1.2264, Acc. No. AEZO02000028; *E. coli* DG131-3, Acc. No AFDV01000051; *E. coli* 9.0111, Acc. No. AEZZ02000022) contained the complete DNA sequence of the SE-PAI. In all the cases the sequences, with a coverage of more than 90%, shared between 97% and 99% of sequence similarity with the one identified in the prototype strain ED 32, indicating a conserved structure for this PAI. Interestingly, the STEC strains fully sequenced were isolated from small ruminants, (*E. coli* 1.2264 and *E. coli* DG131-3), and from a human case of infection (*E. coli* 9.0111). The alignment has been also conducted against the whole sequence of the pO113 plasmid from the STEC O113:H21 strain EH 41 (Acc. No. NC_007365) and involved the entire SE-PAI and the DNA region downstream the *subAB*$_2$

genes, corresponding to the *E. coli yjhATS* operon. The analysis showed that no DNA regions from the PAI or the flanking region were present in the plasmid harbouring the prototype *subAB*$_1$ allele, but the SubAB-coding region itself with a 90% sequence similarity, confirming how previously reported.

***Comparative analysis of the activity of the two SubAB variants***

The Vero cell assays, conducted to evaluate possible differences in the activity of the two SubAB variants, showed that the Subtilase cytotoxin produced by both the alleles was able to induce the CPE up to a dilution of 1:128. This observation is in line with how previously described for the product of *subAB*$_2$ allele. Comparable results were obtained for the two SubAB variants also when the supernatant of the K12 strains containing the *subAB* recombinant plasmids was used in the VCA. In this latter case, however, the effect on the cells was visible only up to a dilution of 1:32, possibly caused by a lower efficiency of expression in the transformants compared to the wild type strains.

***Presence of the subAB genes in human pathogenic E. coli***

The presence of *subAB* genes was investigated in a panel of 107 *E. coli* strains belonging to six different pathotypes and isolated from cases of human disease (Table 3). The association of *subAB* with LEE-negative STEC has previously been assessed in several investigations [10, 22, 24], therefore only STEC strains lacking the *eae* gene were selected for this study. The screening was carried out by PCR using a set of oligonucleotides (RTsubABF/RTsubABR, Table 1) able to prime the amplification of both the *subAB* allelic variants and revealed the presence of *subAB* genes in nine out of the 15 STEC (60%) and one of the 32 ETEC (3%) strains tested. All the strains belonging to the other pathotype were negative. The *subAB*-positive STEC belonged to eight of the ten serogroups tested

(Table 3), and all possessed the *subAB*$_2$ variant, with the exception of one strain which did not react with the primer pairs specific for the two *subAB* variants. No association between the presence of specific *stx*-gene types and the presence of *subAB* could be established (data not shown).

***Detection and characterization of the subAB locus in LEE-negative STEC from human cases of diarrhoea***

On the basis of the results obtained in the preliminary screening, 162 additional LEE-negative STEC from human cases of diarrhoea and belonging to the ten serogroups previously tested were selected and assayed by PCR for the presence of the two allelic variants of *subAB, tia* and, *saa* (Table 4). The *subAB* genes were detected in about 72% of the strains, and the *subAB*$_2$ variant was present in 98.3% of the positive isolates. The *subAB*$_2$ genes were always associated with *tia*, with the exception of five O91 strains. An O181 strain was positive for both *subAB* and *tia*, but did not react in the primers specific for the two *subAB* variants, suggesting that additional *subAB* alleles may exist. Finally, an O91 strain was positive for both the *subAB*$_1$ and *subAB*$_2$ variants, as well as for the *saa* and *tia* genes, suggesting the concomitant presence of the SE-PAI and a plasmid similar to that described in the prototype O113 strain 98NK2. All the other strains were negative for *saa.*

PCR amplification of *tia* yielded a 600 bp product in the majority of the positive isolates, as expected on the basis of the published sequence of the reference ETEC strain IAI39 (GenBank Acc. No. U20318). A 1.8 kb PCR fragment was observed for 21 strains (19 O128, one O91 and one O146). This amplicon size is consistent with the presence of a 1.2 kb IS*2* element interrupting the *tia* sequence*,* as it has been previously described in strain ED 591 (GenBank Acc. No. FJ664545). Two additional strains (one O91 and one O146)

yielded both the 600 bp and 1.8 kb amplicons, suggesting the presence of two copies of *tia*, one of them possibly interrupted by the IS*2* element.

***Detection and characterization of the subAB locus in LEE-negative STEC from healthy sheep***

The high prevalence of the *subAB*$_2$ variant in LEE-negative STEC isolated from patients with diarrhoea prompted us to investigate on the possible animal reservoir of such *E. coli* strains. Since the presence of the *subAB*$_2$ allele has been previously reported in small ruminants, a panel of 123 LEE-negative STEC isolated from sheep and belonging to 10 different serogroups was examined for the presence of the two *subAB* variants, as well as for *tia* and *saa* genes. The presence of *subAB* genes was detected in 108 strains (87.8%), all but two possessing the *subAB*$_2$ variant (Table 5). The latter two strains belonged to serogroups O6 and O123 and did not react with the primer pairs specific for the two *subAB* variants, further supporting the existence of additional *subAB* alleles.

As observed for the human strains, most of the *subAB*$_2$-positive ovine isolates (81.5%) were also positive for *tia.* The 19 *tia*-negative strains belonged to serogroups O166 (14 strains), O91 (3 strains), O76 (1 strain) and O176 (1 strain). None of the strains was positive in the *saa*-specific PCR assay.

## Discussion

Since their first description, SubAB-coding genes have been mainly detected in LEE-negative STEC [8-18], with the exception of two *E. coli* strains that did not produce Stx [22]. In the latter strains, the subtilase was encoded by an allelic variant of the *subAB* genes, termed *subAB*$_2$, which shared 90% homology with the prototype genes, *subAB*$_1$, identified in the virulence plasmid of the LEE-negative O113 STEC strain 98NK2 [1]. The *subAB*$_2$ operon was located on the chromosome, in a putative PAI that contained also the

gene *tia*, encoding an invasion determinant described in ETEC [23]. In the present study, we completed the sequence of this putative PAI, that we termed SE-PAI, and showed that it is constituted by an 8 kb DNA region inserted downstream the *pheV*-tRNA locus in the chromosome of strain ED 32. Beside *subAB*$_2$ and *tia,* SE-PAI carries another virulence gene, the *Shigella flexneri* gene *shiA*, whose product has been described to attenuate the host inflammatory response induced by *Shigella flexneri* infections [26]. The observation that SE-PAI is inserted close to the *pheV*-tRNA locus opens the way to speculations on the evolution of *subAB-positive* STEC strains. This locus represents a hot-spot for the integration of PAIs in different bacterial pathogens, such as the *she* PAI of *S. flexneri* 2a [27] and PAI II$_{AL862}$ of the extraintestinal pathogenic *E. coli* strain AL862 [28]. Moreover, this locus represents the insertion site of the LEE PAI in many STEC strains, including those belonging to serotype O103:H2, one of those most commonly associated with severe disease in humans [29]. It could be hypothesized that the LEE-negative, SE-PAI-positive STEC may have arisen from a mutual exclusion event between this PAI and the LEE, which might have competed for the same integration site. This hypothesis could explain the strong association of the *subAB*$_2$ operon with LEE-negative STEC and its apparent absence in LEE-positive strains.

Thus far, little was known about the presence of the two *subAB* allelic variants among STEC. Recently, a study conducted on different animal species reported a high prevalence of *subAB*$_2$ among LEE-negative strains isolated from small ruminants [14]. In that study, the discrimination between the two allelic variants was based on the intensity of the bands obtained by a non allele-specific PCR amplification of the *subAB* genes, with more intense bands identifying *subAB*$_1$ and faint bands suggesting the presence of *subAB*$_2$ [14]. Another recent study reported that 12 *subAB*-positive O128:H2 STEC strains from different sources were all positive for *tia* [16], suggesting the presence of the *subAB*$_2$-harboring SE-PAI

We adopted a PCR-based strategy for the specific detection of the two allelic variants of *subAB* to investigate their presence in a vast collection of *E. coli* strains isolated from human and animal sources. The screening of a panel of strains isolated from cases of human disease and belonging to different pathotypes confirmed that *subAB* genes are mainly confined to STEC lacking the LEE locus. A further screening of a larger collection of LEE-negative STEC strains isolated from patients with diarrhoea showed that the *subAB* genes were present in more than 70% of the isolates, thus supporting the hypothesis that SubAB may play a role in the disease caused by these STEC strains [1]. Interestingly, almost all the *subAB*-positive strains were positive for the *subAB*$_2$ variant, and the concomitant presence of *tia* suggests that the Subtilase-coding genes were carried by SE-PAI.

The high prevalence of the *subAB*$_2$ variant observed in human LEE-negative STEC prompted us to investigate on the possible sources of such pathogenic *E. coli*. Since the association of *subAB*$_2$ with STEC from small ruminants has been previously suggested [14], we examined a large collection of LEE-negative STEC isolated from sheep by using the same allele-specific PCR strategy. The presence of *subAB* genes was observed in 86% of the strains and, similarly to the human isolates, all the strains possessed the *subAB*$_2$ allelic variant with the exception of two untypeable strains. Again, most *subAB*$_2$-positive strains (84%) were also positive for *tia* suggesting the presence of SE-PAI. Altogether, these observations indicate that *subAB*$_2$ represents the prevalent allelic variant in the SubAB-positive, LEE-negative STEC either colonizing sheep or causing diarrhoea in humans, at least among the strains circulating in European countries.

The *subAB*$_2$ operon was also present in one ETEC strain. This finding may be in relationship with the presence of *tia* in SE-PAI. As a matter of fact, *tia gene* was firstly described as a virulence gene of ETEC [23] suggesting that SE-PAI may have risen as a

mosaic structure through recombination events involving ETEC, before being stably acquired by LEE-negative STEC.

The high prevalence of *subAB*$_2$ genes in STEC of ovine origin suggests that sheep may represent an important animal reservoir of *subAB*$_2$-positive, LEE-negative STEC and confirms the zoonotic origin of these human infections. Further studies are needed to assess: i) the role of other ruminant species in the epidemiology of these infections; ii) to elucidate if *tia* and/or other genes carried by SE-PAI may be involved in the colonization of the host intestinal mucosa by *subAB*$_2$–positive STEC, acting as an alternative to the attaching and effacing machinery.

## References


1. Paton AW, Srimanote P, Talbot UM, Wang H, Paton JC. A new family of potent $AB_5$ cytotoxins produced by Shiga toxigenic *Escherichia coli*. *J Exp Med* 2004; 200: 35-46.
2. Chong DC, Paton JC, Thorpe CM, Paton AW. Clathrin-dependent trafficking of subtilase cytotoxin, a novel $AB_5$ toxin that targets the endoplasmic reticulum chaperone BiP. *Cell Microbiol* 2008; 10: 795-806.
3. Paton AW, Beddoe T, Thorpe CM, Whisstock JC, Wilce MC, Rossjohn J, Talbot UM, Paton JC. $AB_5$ subtilase cytotoxin inactivates the endoplasmic reticulum chaperone BiP. *Nature* 2006; 443: 548-552.
4. Morinaga N, Yahiro K, Matsuura G, Moss J, Noda M. Subtilase cytotoxin, produced by Shiga-toxigenic *Escherichia coli*, transiently inhibits protein synthesis of Vero cells via degradation of BiP and induces cell cycle arrest at G1 by downregulation of cyclin D1. *Cell Microbiol* 2008; 10: 921-929.
5. Yahiro K, Satoh M, Morinaga N, Tsutsuki H, Ogura K, Nagasawa S, Nomura F, Moss J, Noda M. Identification of subtilase cytotoxin (SubAB) receptors whose signalling, in association with SubAB-induced BiP cleavage, is responsible for apoptosis in HeLa cells. *Infect Immun* 2011; 79: 617-27.
6. Yahiro K, Tsutsuki H, Ogura K, Nagasawa S, Moss J, Noda M. Regulation of subtilase cytotoxin-induced cell death by an RNA-dependent protein kinase-like endoplasmic reticulum kinase-dependent proteasome pathway in HeLa cells. *Infect Immun* 2012; 80: 1803-1814.
7. Wolfson JJ, May KL, Thorpe CM, Jandhyala DM, Paton JC, Paton AW. Subtilase cytotoxin activates PERK, IRE1 and ATF6 endoplasmic reticulum stress-signalling pathways. *Cell Microbiol* 2008; 10: 1775-1786.

8. Bugarel M, Beutin L, Martin A, Gill A, Fach P. Micro-array for the identification of Shiga toxin-producing *Escherichia coli* (STEC) seropathotypes associated with Hemorrhagic Colitis and Hemolytic Uremic Syndrome in humans. *Int J Food Microbiol* 2010; 142: 318-329.
9. Buvens G, Lauwers S, Piérard D. Prevalence of subtilase cytotoxin in verocytotoxin-producing *Escherichia coli* isolated from humans and raw meats in Belgium. *Eur J Clin Microbiol Infect Dis* 2010; 29: 1395-1399.
10. Cergole-Novella MC, Nishimura LS, Dos Santos LF, Irino K, Vaz TM, Bergamini AM, Guth BE. Distribution of virulence profiles related to new toxins and putative adhesins in Shiga toxin-producing *Escherichia coli* isolated from diverse sources in Brazil. *FEMS Microbiol Lett* 2007; 274: 329-334.
11. Irino K, Vieira MA, Gomes TA, Guth BE, Naves ZV, Oliveira MG, dos Santos LF, Guirro M, Timm CD, Pigatto CP, Farah SM, Vaz TM. Subtilase cytotoxin-encoding *subAB* operon found exclusively among Shiga toxin-producing *Escherichia coli* strains. *J Clin Microbiol* 2010; 48: 988-990.
12. Karama M, Johnson RP, Holtslander R, McEwen SA, Gyles CL. Prevalence and characterization of verotoxin-producing *Escherichia coli* (VTEC) in cattle from an Ontario abattoir. *Can J Vet Res* 2008; 72: 297–302.
13. Khaitan A, Jandhyala DM, Thorpe CM, Ritchie JM, Paton AW. The operon encoding SubAB, a novel cytotoxin, is present in shiga toxin-producing *Escherichia coli* isolates from the United States. *J Clin Microbiol* 2007; 45: 1374-1375.
14. Orden JA, Horcajo P, de la Fuente R, Ruiz-Santa-Quiteria JA, Domínguez-Bernal G, Carrión J. Subtilase cytotoxin-coding genes in verotoxin-producing *Escherichia coli* strains from sheep and goats differ from those from cattle. *Appl Environ Microbiol* 2011; 77: 8259-8264.

**15.** Paton AW, Paton JC. Multiplex PCR for direct detection of Shiga toxigenic *Escherichia coli* strains producing the novel subtilase cytotoxin. *J Clin Microbiol* 2005; 43: 2944-2947.

**16.** Sánchez S, Beristain X, Martínez R, García A, Martín C, Vidal D, Díaz-Sánchez S, Rey J, Alonso JM, Herrera-León S. Subtilase cytotoxin encoding genes are present in human, sheep and deer intimin-negative, Shiga toxin-producing *Escherichia coli* O128:H2. *Vet Microbiol* 2012; 159: 531-535.

**17.** Wolfson JJ, Jandhyala DM, Gorczyca LA, Qadeer Z, Manning SD, Hadler J, Rudrik JT, Thorpe CM. Prevalence of the operon encoding subtilase cytotoxin in non-O157 Shiga toxin-producing *Escherichia coli* isolated from humans in the United States. *J Clin Microbiol* 2009; 47: 3058-3059.

**18.** Wu Y, Hinenoya A, Taguchi T, Nagita A, Shima K, Tsukamoto T, Sugimoto N, Asakura M, Yamasaki S. Distribution of virulence genes related to adhesins and toxins in shiga toxin producing *Escherichia coli* strains isolated from healthy cattle and diarrheal patients in Japan. *J Vet Med Sci* 2010; 72: 589-597.

**19.** Frankel G, Phillips AD, Rosenshine I, Dougan G, Kaper JB, Knutton S. Enteropathogenic and enterohaemorrhagic *Escherichia coli*: more subversive elements. *Mol Microbiol* 1998; 30: 911-921.

**20.** Wang H, Paton JC, Paton AW. Pathologic changes in mice induced by subtilase cytotoxin, a potent new *Escherichia coli* $AB_5$ toxin that targets the endoplasmic reticulum. *J Infect Dis* 2007; 196: 1093-2101.

**21.** Paton AW, Srimanote P, Woodrow MC, Paton JC. Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. *Infect Immun* 2001; 69: 6999-7009.

**22.** Tozzoli R, Caprioli A, Cappannella S, Michelacci V, Marziano ML, Morabito S. Production of the subtilase $AB_5$ cytotoxin by Shiga toxin-negative *Escherichia coli*. *J Clin Microbiol* 2010; 48: 178-183.

**23.** Fleckenstein JM, Kopecko DJ, Warren RL, Elsinghorst EA. Molecular characterization of the *tia* invasion locus from enterotoxigenic *Escherichia coli*. *Infect Immun* 1996; 64: 2256-2265.

**24.** Franke J, Franke S, Schmidt H, Schwarzkopf A, Wieler LH, Baljer G, Beutin L, Karch H. Nucleotide sequence analysis of enteropathogenic *Escherichia coli* (EPEC) adherence factor probe and development of PCR for rapid detection of EPEC harboring virulence plasmids. J Clin Microbiol 1994 32:2460-2463.

**25.** Paton AW, Paton JC. Direct detection and characterization of Shiga toxigenic *Escherichia coli* by multiplex PCR for *stx1*, *stx2*, *eae*, *ehxA*, and *saa*. *J Clin Microbiol* 2002; 40: 271-274.

**26.** Ingersoll MA, Zychlinsky A. ShiA abrogates the innate T-cell response to *Shigella flexneri* infection. *Infect Immun* 2006; 74: 2317-2327.

**27.** Al-Hasani K, Rajakumar K, Bulach D, Robins-Browne R, Adler B, Sakellaris H. Genetic organization of the *she* pathogenicity island in *Shigella flexneri 2a*. *Microb Pathog* 2001; 30: 1-8.

**28.** Lalioui L, Le Bouguénec C. *afa-8* Gene cluster is carried by a pathogenicity island inserted into the tRNA(Phe) of human and bovine pathogenic *Escherichia coli* isolates. *Infect Immun* 2001; 69: 937-948.

**29.** Jores J, Rumer L, Kiessling S, Kaper JB, Wieler LH. A novel locus of enterocyte effacement (LEE) pathogenicity island inserted at *pheV* in bovine Shiga toxin-producing *Escherichia coli* strain O103:H2. *FEMS Microbiol Lett* 2001; 204: 75-79.

TABLE 1. PCR primers used in this study for the characterization of the SubAB–encoding PAI, for the detection of *subAB*, *tia*, *saa*, and for identifying the two allelic variants of the *subAB* operon.

| Primer name | *Gene target* | Sequence | GenBank accession number (nucleotide position) | Reference |
|---|---|---|---|---|
| RTsubABF | *subA* | 5'-GCAGATAAATACCCTTCACTTG-3' | AF399919 (13856-13835) | 1 |
| RTsubABR | *subB* | 5'-ATCACCAGTCCACTCAGCC-3' | AF399919 (13625-13643) | 1 |
| tia lo | *tia* | 5'-TCCATGCGAAGTTGTTATCA-3' | U20318 (577-558) | 22 |
| tia up | *tia* | 5'-GAAATGAAAAAGATTATTGCGG-3' | U20318 (7-28) | 22 |
| SubAF | *subA* | 5'-GTACGGACTAACAGGGAACTG-3' | AF399919 (14944-14964) | 1 |
| subA_anchor | *subA* | 5'-CATGACGTGAGGCAATGAGT-3' | JQ994271 (6565-6545) | This study |
| tRNA_phe_anchor | *tRNAphe* | 5'-ATTGAAAATCCCCGTGTCCT-3' | CU928164 (4802156- 4802137) | This study |
| subA_startF | *subA* | 5'-CCCTGTAACATATTGACCAGCA-3' | JQ994271 (6208-6230) | This study |
| subB_end_F | *subB* | 5'-GGCCTTCGTTTTCCACATTA-3' | JQ994271 (7550-7570) | This study |
| subB_end_R | *subB* | 5'-TTATGAGTTCTTTTTCCTGTCAGG-3' | AF399919 (13283-13306) | 22 |
| saa up | *saa* | 5'-CGTGATGAACAGGCTATTGC-3' | AF399919 (7712- 7731) | 25 |
| saa lo | *saa* | 5'-ATGGACATGCCTGTGGCAAC-3' | AF399919 (7830- 7811) | 25 |

TABLE 2. Summary of ORFs in SE-PAI of ED 32 strain identified by significant similarity (BLASTP search).

| ORF# | Location in SE-PAI | Denomination in SE-PAI | Closest informative protein match | Homologues Acc. No. | Aminoacidic (AA) identities | Positive AA substitutions |
|---|---|---|---|---|---|---|
| 1 | 144-1328 | integrase | site-specific recombinase, phage integrase family [*Escherichia coli* 99.0741] | ZP_11999107 | 394/394 (100%) | 394/394 (100%) |
| 2 | 1627-2700 | *shiA* | putative homolog to shiA (SHI-2 pathogenicity island of *Shigella flexneri*) [*Escherichia coli* IAI39] | CAR19509 | 320/353 (91%) | 332/353 (94%) |
| 3 | 3131-4438 | sulfatase | sulfatase family protein [*Escherichia coli* STEC_DG131-3] | ZP_12260098 | 431/435 (99%) | 432/435 (99%) |
| 4 | 4726-5472 | *tia* | tia invasion determinant [*Escherichia coli* 1.2264] | ZP_11990814 | 247/248 (99%) | 248/248 (100%) |
| 5 | 6262-7305 | *subA* | subtilase family protein [*Escherichia coli* STEC_EH250] | EGW88118 | 347/347 (100%) | 347/347 (100%) |
| 6 | 7322-7744 | *subB* | subtilase cytotoxin subunit B [*Escherichia coli*] | ACV40235 | 140/140 (100%) | 140/140 (100%) |

TABLE 3. Results of the PCR analyses for the presence of *subAB* in human *E. coli* strains belonging to different pathotypes (tEPEC: typical EPEC; aEPEC: atypical EPEC)

| *E. coli* pathotype (No. of strains) | Serogroup (No. of strains) | n. of subAB-positive strains (allelic variant) |
|---|---|---|
| STEC (15) | O76 (1) | 1 (*subAB*$_2$) |
| | O78 (1) | 1 (*subAB*$_2$) |
| | O91 (2) | 1 (*subAB*$_2$) |
| | O113 (1) | 1 (*subAB*$_2$) |
| | O117 (2) | 0 |
| | O128 (2) | 2 (*subAB*$_2$) |
| | O146 (2) | 1 (*subAB*$_2$) |
| | O174 (1) | 0 |
| | O181 (1) | 1 (untypeable) |
| | Orough (2) | 1 (*subAB*$_2$) |
| tEPEC (3) | O119 (2) | 0 |
| | O145 (1) | 0 |
| aEPEC (23) | O8 (1) | 0 |
| | O26 (2) | 0 |
| | O55 (2) | 0 |
| | O103 (3) | 0 |
| | O111 (1) | 0 |
| | O114 (1) | 0 |
| | O119 (1) | 0 |
| | O125 (1) | 0 |
| | O127 (1) | 0 |
| | O128 (3) | 0 |
| | O145 (3) | 0 |
| | O157 (4) | 0 |
| ETEC (32) | O n.t.* (1) | 0 |
| | O6 (3) | 0 |
| | O8 (3) | 0 |
| | O25 (4) | 0 |
| | O44 (1) | 0 |
| | O153 (4) | 0 |
| | O164 (1) | 0 |
| | Orough (15) | 1 (*subAB*$_2$) |
| EAEC (20) | O25 (1) | 0 |
| | O92 (2) | 0 |
| | O+** (17) | 0 |
| EIEC (13) | O121 (1) | 0 |
| | O n.t.* (12) | 0 |
| UPEC (1) | O102 (1) | 0 |

* n.t.: not tested

** O+: the O antigen is present but couldn't be typed

TABLE 4. Results of the PCR analyses for the presence of the *subAB* allelic variants, *tia*, and *saa* in in LEE-negative STEC strains of human origin.

| Genotype | | | | No. of strains positive for | | | | |
|---|---|---|---|---|---|---|---|---|
| **Serotype** | ***stx1*** | ***stx2*** | **No. of strains** | ***subAB*** | ***subAB$_2$*** | ***tia*** | ***subAB$_1$*** | ***saa*** |
| O76:H19 | + | - | 10 | 10 | 10 | 10 | 0 | 0 |
| O78:H- | + | - | 3 | 3 | 3 | 3 | 0 | 0 |
| O78:H- | + | + | 1 | 1 | 1 | 1 | 0 | 0 |
| O91:H- | + | - | 11 | 1 | 1 | 0 | 0 | 0 |
| O91:H- | - | + | 1 | 1 | 1 | 1 | 0 | 0 |
| O91:H- | + | + | 30 | 30 | 29 | 25 | 1* | 1* |
| O113:H4 | + | + | 6 | 5 | 5 | 5 | 0 | 0 |
| O117:H7 | + | - | 15 | 0 | - | - | - | - |
| O117:H- | + | - | 4 | 0 | - | - | - | - |
| O128:H2 | - | + | 27 | 26 | 26 | 26 | 0 | 0 |
| O128:H2 | + | + | 4 | 4 | 4 | 4 | 0 | 0 |
| O146:H21 | + | + | 26 | 24 | 24 | 24 | 0 | 0 |
| O146:H28 | - | + | 5 | 2 | 2 | 2 | 0 | 0 |
| O146:H28 | + | + | 2 | 2 | 2 | 2 | 0 | 0 |
| O174:H21 | - | + | 4** | 0 | - | - | - | - |
| O181:H16 | + | - | 3 | 3 | 2 | 3 | 0 | 0 |
| Orough:H- | + | - | 6 | 1 | 1 | 1 | 0 | 0 |
| Orough:H- | + | + | 4 | 4 | 4 | 4 | 0 | 0 |
| ***Total*** | | | ***162*** | ***117*** | ***115*** | ***111*** | ***1*** | ***1*** |

*: One single strain was positive for *subAB$_1$*, *subAB$_2$*, *saa* and *tia* simultaneously.

**: The strains have not been tested for the *subAB* variants, *tia* and *saa*, being negative to the *subAB* screening assay (see text for details)

TABLE 5. Results of the PCR analyses for the presence of the *subAB* allelic variants, *tia*, and *saa* in in LEE-negative STEC strains isolated from healthy sheep.

| | | No. of strains positive for | | | | |
|---|---|---|---|---|---|---|
| **Serotype** | **No. of strains** | ***subAB*** | ***subAB*$_2$** | ***tia*** | ***subAB*$_1$** | ***saa*** |
| O5 | 11 | 8 | 8 | 8 | 0 | 0 |
| O6 | 7 | 5 | 4 | 5 | 0 | 0 |
| O76 | 12 | 12 | 12 | 11 | 0 | 0 |
| O87 | 10* | 0 | - | - | - | - |
| O91 | 26 | 26 | 26 | 23 | 0 | 0 |
| O123 | 5 | 5 | 4 | 5 | 0 | 0 |
| O128 | 6 | 6 | 6 | 6 | 0 | 0 |
| O146 | 19 | 19 | 19 | 19 | 0 | 0 |
| O166 | 14 | 14 | 14 | 0 | 0 | 0 |
| O176 | 13 | 13 | 13 | 12 | 0 | 0 |
| ***Total*** | ***123*** | ***108*** | ***106*** | ***89*** | ***0*** | ***0*** |

*: The strains have not been tested for the *subAB* variants, *tia* and *saa*, being negative to the *subAB* screening assay (see text for details)

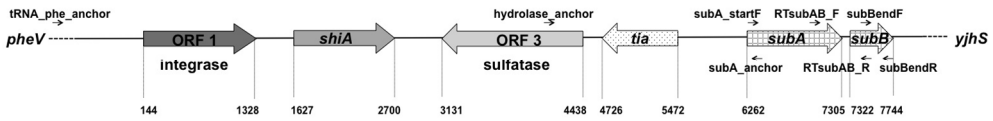



FIGURE 1. Map of the Subtilase-Encoding Pathogenicity Island (SE-PAI), spanning the 8058 bp comprised between the pheV
tRNA locus and the E. coli gene yjhS. The ORFs and the annotated genes, together their positions in the PAI sequence are shown (GenBank Acc. No. JQ994271). The locations of the primers used to characterize SE-PAI are also indicated.
549x63mm (72 x 72 DPI)

# 4. Identification of a virulence-associated MGE of VTEC serogroups associated with severe human disease (OI-57)

# 4. Identification of a virulence-associated MGE of VTEC serogroups associated with severe human disease (OI-57)

*Introduction*

VTEC are a heterogeneous pathogroup of *E. coli* whose pathogenicity relies on the production of the verocytotoxins (VT) and on the presence of accessory virulence genes acquired by horizontal gene transfer, which have not been completely identified yet. VTEC have been categorized into seropathotypes (SPTs) based on their phenotypic and molecular characteristics and the clinical features of the associated diseases. SPTs range from A to E, according to a decreasing rank of pathogenicity.

The purpose of the paper presented in this chapter was the identification of the genetic determinants conferring virulence to the most pathogenic VTEC and the analysis of their distribution in VTEC strains belonging to different SPTs.

*Results*

Microarrays have been used to compare the whole genomes of VTEC strains belonging to SPTs B (13 strains), C (four strains), and D (one strain) with that of VTEC O157 (SPT A). Total DNA from 18 selected strains was hybridized onto commercial microarray slides containing the complete gene set of the *E. coli* O157:H7 strains EDL933 and Sakai (belonging to SPT A) and *E. coli* K-12 strain MG1655 as a control strain. We identified 87 open reading frames (ORFs), which were absent in the SPTs C and D, but gave strong hybridization signals with the SPT B strains, which were selected for further investigation. Sixty-one of these ORFs were part of metabolic pathways or part of Genomic Islands (GI) also found in strains belonging to SPTs C and D and were not furthered considered. The remaining 26 ORFs were investigated in a larger panel of *E. coli* strains including 61 VTEC isolates and 27 non-pathogenic *E. coli*. Eight of these ORFs were found to be associated with SPTs A and B, with four of them being part of the same GI, termed OI-57 in VTEC O157 reference strain EDL933.

Given the association of this GI with the most pathogenic VTEC SPTs, we investigated on its integrity by a PCR screening targeting six ORFs scattered along its whole length and demonstrated that the entire OI-57 is present in 94% of the strains belonging to SPTs A and B and none of the strains belonging to other SPTs, suggesting that OI-57 is part of the virulome of the VTEC causing severe disease in humans. (Publication: *Infection and Immunity*, 78(11):4697-4704. Impact Factor 4.165).

*Conclusions*

The genomic island termed OI-57 has been shown to be consistently present in VTEC strains belonging to SPTs A and B but not in the less pathogenic SPTs C and D. OI-57 contains putative virulence-associated genes such as *adfO*, which encodes a factor that has been shown to promote the adherence of VTEC O157 to *in vitro*-cultured cells, and *ckf*, encoding a putative phage-associated killer protein.

Through an extensive PCR screening of VTEC strains, OI-57 was demonstrated to be entire and stable in the positive clonal lineages. OI-57 was also present in a high proportion of the human enteropathogenic *E. coli* genomes assayed, suggesting that it could be involved in the *attaching-and-effacing* colonization of the intestinal mucosa. OI-57 thus appears as a pathogenicity island part of the virulome of VTEC causing severe disease in humans. Further studies are needed to elucidate its role in the pathogenesis of VTEC infections.

INFECTION AND IMMUNITY, Nov. 2010, p. 4697–4704 Vol. 78, No. 11
0019-9567/10/$12.00 doi:10.1128/IAI.00512-10


# OI-57, a Genomic Island of *Escherichia coli* O157, Is Present in Other Seropathotypes of Shiga Toxin-Producing *E. coli* Associated with Severe Human Disease[▽]

Lejla Imamovic,† Rosangela Tozzoli, Valeria Michelacci, Fabio Minelli, Maria Luisa Marziano, Alfredo Caprioli, and Stefano Morabito*

*European Union Reference Laboratory for Escherichia coli, Istituto Superiore di Sanità, Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare, Rome, Italy*

Received 17 May 2010/Returned for modification 11 June 2010/Accepted 27 August 2010

**Strains of Shiga toxin-producing *Escherichia coli* (STEC) are a heterogeneous *E. coli* group that may cause severe disease in humans. STEC have been categorized into seropathotypes (SPTs) based on their phenotypic and molecular characteristics and the clinical features of the associated diseases. SPTs range from A to E, according to a decreasing rank of pathogenicity. To define the virulence gene asset ("virulome") characterizing the highly pathogenic SPTs, we used microarray hybridization to compare the whole genomes of STEC belonging to SPTs B, C, and D with that of STEC O157 (SPT A). The presence of the open reading frames (ORFs) associated with SPTs A and B was subsequently investigated by PCR in a larger panel of STEC and in other *E. coli* strains. A genomic island termed OI-57 was present in SPTs A and B but not in the other SPTs. OI-57 harbors the putative virulence gene *adfO*, encoding a factor enhancing the adhesivity of STEC O157, and *ckf*, encoding a putative killing factor for the bacterial cell. PCR analyses showed that OI-57 was present in its entirety in the majority of the STEC genomes examined, indicating that it represents a stable acquisition of the positive clonal lineages. OI-57 was also present in a high proportion of the human enteropathogenic *E. coli* genomes assayed, suggesting that it could be involved in the attaching-and-effacing colonization of the intestinal mucosa. In conclusion, OI-57 appears to be part of the virulome of pathogenic STEC and further studies are needed to elucidate its role in the pathogenesis of STEC infections.**

Shiga toxin (Stx)-producing *Escherichia coli* (STEC) strains cause severe human diseases such as hemorrhagic colitis and the life-threatening hemolytic uremic syndrome (HUS) (20, 32). STEC strains are zoonotic pathogens, and ruminants, particularly cattle, are recognized as their main natural reservoir (3). STEC strains may belong to a large number of serotypes, but only some of them have been firmly associated with severe human infections (20, 36). *E. coli* O157:H7 causes most of the cases of severe disease worldwide, but infections sustained by STEC strains belonging to serogroups other than O157, such as O26, O111, O103, O145 and O121, have been increasingly reported (2, 3, 4, 26, 33, 36, 39).

The pathogenicity of STEC mainly relies on the presence of virulence genes located on mobile genetic elements (MGEs). These include the lambdoid bacteriophages transducing the Stx-coding genes (22) and the large virulence plasmid carrying the genes governing the production of enterohemolysin (28) and other putative virulence factors. In STEC O157, this plasmid also contains *toxB*, a gene that encodes a protein likely involved in the colonization of the gut mucosa (30, 34). Other important virulence properties are carried by pathogenicity islands (PAIs) such as the "locus of enterocyte effacement" (LEE) (16), which governs the capability of STEC to colonize the intestinal mucosa of the host with a peculiar mechanism known as attaching-and-effacing (A/E) (16, 20). Another PAI, termed O island 122 (OI-122) in the sequence of STEC O157 strain EDL933 (24), has been frequently found in STEC strains usually associated with severe human disease (9, 11, 19). This PAI carries the large virulence gene *efa1-lifA*, which appears to be involved in the colonization of the intestinal mucosa and in the inhibition of the host immune response (10, 21, 31). Beside these MGEs, many horizontally acquired DNA regions carrying genes encoding putative virulence factors or proteins of unknown function have been described in the genome of the STEC strains fully or partially sequenced so far. More than 170 of these regions have been annotated as genomic islands (GIs) in the sequences of the STEC O157 strains EDL933 and RIMD0509952/VT2 Sakai (7, 25).

Due to the high genomic plasticity of *E. coli*, MGEs can be easily exchanged among strains or lost, generating STEC clones with different genetic assets and, as a consequence, different degrees of virulence. Karmali et al. proposed to group STEC strains into five seropathotypes (SPTs), from A through E, according to their reported frequencies in human illness, their known association with outbreaks and severe disease, and the presence of MGEs such as the LEE and the OI-122 PAI (9). SPT A includes STEC strains of serotypes O157:H7 and O157:NM, associated with both outbreaks and sporadic cases of HUS. SPT B includes STEC strains that are capable of causing HUS and outbreaks but less commonly than STEC O157, such as those belonging to serotypes O26:H11, O103:H2, O111:NM, O121:H19, and O145:NM, whereas SPT C in-

* Corresponding author. Mailing address: Dipartimento di Sanità, Pubblica Veterinaria e Sicurezza Alimentare, Istituto Superiore di Sanità, Viale Regina Elena, 299, 00161 Rome, Italy. Phone: 39-06-49903081. Fax: 39-06-49387077. E-mail: stefano.morabito@iss.it.

† Present address: Department of Microbiology, Faculty of Biology, University of Barcelona, Barcelona, Spain.

▽ Published ahead of print on 7 September 2010.

cludes strains belonging to serotypes O91:H21 and O113:H21, which lack the LEE and have sometimes been associated with HUS (9). STEC serotypes associated with diarrhea but not with HUS, and STEC strains circulating only in animal reservoirs, have been included in SPTs D and E, respectively (9). This classification scheme takes into consideration the clinical aspects of the disease as well as the molecular features reported to play a role in STEC pathogenicity. However, the molecular bases of STEC infections have not been completely understood and the whole genetic asset conferring full virulence to SPTs A and B has not yet been completely identified.

In an attempt to clarify this issue, we compared the entire genome of STEC O157 with those of STEC strains belonging to SPTs B, C, and D. This approach allowed us to identify a putative PAI (termed OI-57) in the sequence of the STEC O157 EDL933 strain (25) that was consistently present in STEC strains belonging to SPTs A and B but not in the less pathogenic SPTs C and D. The association of OI-57 with SPTs A and B was confirmed by examining a panel of STEC strains that included isolates belonging to SPTs from type A to type E as well as nonpathogenic *E. coli* strains.

## MATERIALS AND METHODS

**Bacterial strains.** All the *E. coli* strains used in this study were part of the culture collection of the Istituto Superiore di Sanità and have been described in reports from previous studies (17, 18, 24, 37, 38). All the strains had been isolated in Italy between 1988 and 2008. STEC O157 reference strains EDL933 and RIMD0509952/VT2 Sakai, together with *E. coli* K-12 strain MG1655, were also included in the study. These strains were kindly provided by Roberto La Ragione, Veterinary Laboratories Agency, Weybridge, United Kingdom.

Eighteen STEC strains were used in the microarray experiments. They belonged to (i) serogroups O26 (five strains), O111 (four strains), O103 (two strains), and O145 (two strains), as representatives of seropathotype B; (ii) serogroups O113 (two strains) and O91 (two strains), as representatives of seropathotype C; and (iii) serogroup O45 (one strain), as a representative of seropathotype D.

The PCR analyses for the validation of the microarray results were performed using a panel of 61 additional STEC strains. SPT A included nine O157 strains. SPT B strains belonged to serogroups O26 (nine strains), O111 (ten), O145 (seven), O103 (four), and O121 (two); SPT C strains belonged to serogroups O113 (nine) and O91 (four). SPTs D and E were represented by seven strains of animal origin that were negative for the presence of the intimin-coding *eae* gene (23) and belonged to serogroups O23, O25, O8, O73, O109, and O nontypeable (ONT).

The presence of the putative virulence gene in OI-57 was also assessed in a panel of Stx-negative *E. coli* isolates belonging to other pathogroups. These included 42 enteropathogenic *E. coli* (EPEC) isolates, nine enterotoxigenic *E. coli* (ETEC) isolates, eight enteroaggregative *E. coli* (EAEC) isolates, and eight cytotoxic necrotizing factor 1 (CNF1)-producing *E. coli* (NTEC) isolates. The EPEC strains were characterized as typical or atypical on the basis of the presence or the absence of the EPEC adherence factor (EAF) plasmid as described by Franke et al. (6). The ETEC strains belonged to serogroups O6, O64, and O147; the EAEC strains belonged to serogroups O86, O111, O126, and O128; and the NTEC strains belonged to serogroups O2, O6, O75, and O83.

Twenty-seven nonpathogenic *E. coli* isolates from the *Escherichia coli* Reference (ECOR) collection (23) and the K-12 LE 392 strain were also examined.

**Microarray hybridizations.** The microarray slides used in this study were purchased from Ocimum Biosolutions Ltd. (Hyderabad, India). They were composed of 6,176 oligonucleotides (50-mer) representing almost all the coding genes of the three control strains. In particular, the slides were spotted with oligonucleotides specific for 4,288 open reading frames (ORFs) of *E. coli* K-12 strain MG1655, 5,358 ORFs of *E. coli* O157:H7 strain RIMD0509952, and 5,336 ORFs of *E. coli* O157:H7 strain EDL933.

Total DNA of the test strains was purified from overnight bacterial cultures by using a genomic DNA extraction kit (Gentra Systems) according to the manufacturer's instructions. DNA was labeled with a BioPrime DNA labeling system (InVitrogen Life Technologies, Carlsbad, CA). For a 50-μl reaction mixture, 2.5 μg of DNA was combined with 15 μg of random octamers, heated to 95°C for 5 min, and chilled in ice. The remaining components were added to achieve the final concentration as follows: 0.12 mM dATP-dGTP-dCTP, 0.06 mM dTTP, 0.02 mM Cy3 or Cy5-dUTP (Amersham Pharmacia Biotech, Little Chalfont, United Kingdom), and 40 units of the Klenow fragment of *E. coli* DNA polymerase. The reaction mixture was incubated at 37°C for 3 h, and the labeled DNA was purified by using a QIAquick PCR purification kit (Qiagen, Chatsworth, CA) according to the manufacturer's instructions. The DNA probe, composed of one of the test DNAs labeled with Cy3 (15 μl) and a mixture of DNA from the three control strains labeled with Cy5 (15 μl), was mixed with 15 μl of hybridization buffer (Ocimum Biosolutions Ltd., Hyderabad, India), heated to 95°C for 5 min, and chilled on ice before hybridization on the slide. The mixture of the three control DNAs was used to check the efficiency of hybridization on the whole slide surface.

Hybridizations were carried out overnight under glass coverslips in a sealed wet box at 42°C. Following hybridization, slides were washed for 2 min in 1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.05% sodium dodecyl sulfate (SDS) at 42°C and two times for 2 min in 0.06× SSC at room temperature. Following washes, the slides were dried by centrifugation in conical 50-ml tubes.

Hybridized slides were analyzed by using a ProScanArray microarray scanner (Perkin Elmer, Downers Grove, IL). Computer analysis of images was conducted by using ScanArray software (Perkin Elmer, Downers Grove, IL) and MeV software (27). An ORF was considered present in the test DNA when the corresponding spot in the hybridized slide showed a mean fluorescence value higher than that of the background plus 2 standard deviations.

**PCR validation of the microarray results.** The major differences identified by the microarray analyses were further characterized by developing PCR assays aimed at assessing the presence of the ORFs that had been detected in SPTs A and B but not in SPTs C and D. PCR primers were deployed by using the primer-BLAST tool available at NCBI website (http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome) and are listed in Table 1.

All the PCRs were performed using 50 ng of total DNA as the template, 1.5 mM $MgCl_2$, 200 μM deoxynucleoside triphosphate, 1 μM each primer, and 1 U of *Taq* polymerase.

**PCR detection and analysis of the OI-57 genomic island.** The presence of the OI-57 GI was investigated using primer pairs (Table 1) corresponding to the ORFs whose products were annotated with a putative function, namely, the putative virulence factor *adfO* and ORF Z2054, encoding a putative bacterial cell killing factor, according to the EDL933 strain sequence (GenBank accession no. AE005174). Six additional primer pairs, corresponding to as many ORFs (Z2066, Z2096, Z2104, Z2105, Z2148, and Z2150), were used to assess the integrity of the OI-57 GI (Table 1 and Fig. 1).

## RESULTS

**Whole-genome comparison between STEC O157 and STEC strains belonging to SPTs B, C, and D.** Microarrays were used to compare the whole genomes of STEC strains belonging to SPTs B (13 strains), C (four strains), and D (one strain) with that of STEC O157 (SPT A). Total DNA from each of the selected 18 strains was hybridized onto commercial microarray slides containing the complete gene set of the *E. coli* O157:H7 strains EDL933 and RIMD0509952/VT2 Sakai and *E. coli* K-12 strain MG1655.

The analysis of the microarray profiles identified 87 ORFs absent in the strains belonging to SPTs C and D but giving strong hybridization signals with the SPT B strains (not shown). These ORFs represented the most evident difference between the groups studied and were selected for further investigation.

Sequence and functional analyses conducted *in silico* showed that some (61 out of 87) of the ORFs identified were part of *E. coli* metabolic pathways, part of the Stx1- or Stx2-converting phages, or part of the K-12-related Rac phage; other ORFs were part of GIs that have also been found in strains belonging to SPTs C and D. None of these ORFs were further investigated, since they were not considered potential markers of the

TABLE 1. PCR primers used in this study[a]

| Primer name | ORF name | ORF product | Primer sequence | Primer location (nt position) |
|---|---|---|---|---|
| Z0953 Fwd | Z0953 | Hypothetical protein | GGCTTTAACGTGCGCTAACT | 897927–897946 |
| Z0953 Rev | | | TCATCTTTTCCTCGTGCATC | 898046–898026 |
| Z0958 Fwd | Z0958 | Unknown protein encoded by CP-933K | ATGGATTTCACCAGCAGGAT | 902731–902750 |
| Z0958 Rev | | | ATTCATTCTGTTCTTTCATCATCG | 902830–902807 |
| Z0989 Fwd | Z0989 | Unknown protein encoded by prophage CP-933K | AGAAACCTGACTTCGCCTGA | 929184–929203 |
| Z0989 Rev | | | CTTTTGTTGCATCCTCAGCA | 929601–929582 |
| Z1142 Fwd | Z1142 | Urease-associated protein UreD | TGTCGGTCTGCTCACTGTTC | 1078712–1078731 |
| Z1142 Rev | | | TCTCCAGACGGTTAGCGAGT | 1079115–1079096 |
| Z1154 Fwd | Z1154 | Putative colicin immunity protein | TCCAGTTTCTGGTGATGTTTTG | 1085954–1085975 |
| Z1154 Rev | | | TGGCCCTTCCTTGTTCAATA | 1086158–1086139 |
| Z1174 Fwd | Z1174 | Putative tellurium resistance protein TerC | TAATGTTCGCGTTTGACTCG | 1103602–1103621 |
| Z1174 Rev | | | CGATGCTGTAACCATGATGC | 1103838–1103819 |
| Z1214 Fwd | Z1214 | Hypothetical protein | CCAGTGAACATATTTCCGGC | 1139766–113978 |
| Z1214 Rev | | | TGGGTGAAATGAATGAGGTTT | 1139897–1139877 |
| Z1348 Fwd | Z1348 | Hypothetical protein | TAAAACTGCGTCTGGCTGTG | 1267374–1267393 |
| Z1348 Rev | | | TTCTTTTTATCACCGGCCAC | 1267514–1267495 |
| Z1368 Fwd | Z1368 | Hypothetical protein | AGAAGAAGTGGTGTCACGGG | 1282029–1282048 |
| Z1368 Rev | | | GTTATCCGCCTTCATGGTGT | 1282191–1282172 |
| Z1385 Fwd | Z1385 | Putative secreted protein encoded by cryptic prophage CP-933 M | TTAGGTGCCGGAGGAATATG | 1294045–1294064 |
| Z1385 Rev | | | AAGACCTTTGCCTGATGTGG | 1294428–1294409 |
| Z1611 Fwd | Z1611 | putative phage inhibition, colicin resistance and tellurite resistance protein | TGGTGTTTTATGGTCAGCCA | 1496984–1497003 |
| Z1611 Rev | | | TATGCTTCCATTCACGTCCA | 1497736–1497717 |
| Z1780 Fwd | Z1780 | Hypothetical protein | CGTGAGGCAGCACAAAACTA | 1634608–1634627 |
| Z1780 Rev | | | GACGCGTAATGCGTGTTCTA | 1634991–1634972 |
| Z1914 Fwd | Z1914 | Putative minor tail fiber protein CP-933X | TTTTGAGATGAACGGGAAGG | 1736664–1736683 |
| Z1914 Rev | | | TAATCTTTGTCTGTCCCCGC | 1736906–1736887 |
| Z2053 Fwd | adfO | Putative intestinal colonization factor | TGG TGG CCC GCA TAC AGC | 1860232–1860249 |
| Z2053 Rev | | | TGC CCA GTC AGC CCA GGT TA | 1860733–1860714 |
| Z2054 Fwd | ckf | Putative killer protein | ATG CTC GTC ACA TAT AGA TTG | 1861195–1861215 |
| Z2054 Rev | | | GTT CGT AAG CTG TGA AGA CA | 1861396–1861377 |
| Z2066 Fwd | Z2066 | Unknown protein encoded by prophage CP-933O | TAAGGCACCGGAAATCAAAC | 1867809–1867828 |
| Z2066 Rev | | | TGCATCAGATTCACTCTGCC | 1868228–1868209 |
| Z2096 Fwd | Z2096 | Unknown protein encoded within prophage CP-933O | TTCTTTGCATGCCTCTCTGA | 1886847–1886828 |
| Z2096 Rev | | | CGGGAACGATGAAGGTAAAA | 1886665–1886684 |
| Z2097 Fwd | Z2097 | Unknown protein encoded within prophage CP-933O | TTCTTTGCATGCCTCTCTGA | 1886847–1886828 |
| Z2097 Rev | | | CGGGAACGATGAAGGTAAAA | 1886665–1886684 |
| Z2098 Fwd | Z2098 | Unknown protein encoded within prophage CP-933O | GCTGAATGATTCCCATGTCA | 1887762–1887743 |
| Z2098 Rev | | | ACCCACGAAAGAGCAGAACA | 1887608–1887627 |
| Z2104 Fwd | Z2104 | Putative ARAC-type regulatory protein | CATGGTTTGCTCCCCAATTA | 1892233–1892252 |
| Z2104 Rev | | | CGGAATGTCAGTCAGACGAG | 1892440–1892421 |
| Z2105 Fwd | Z2105 | Unknown protein encoded within prophage CP-933O | CTTCAAAAAGAAAACCCGCA | 1893385–1893404 |
| Z2105 Rev | | | GATCTTCGCTTTCTCGGATG | 1893523–1893504 |
| Z2121 Fwd | Z2121 | Unknown protein encoded within prophage CP-933O | CAAGCAGCATTCTTGCAGAC | 1904821–1904840 |
| Z2121 Rev | | | GCTGACATGCGTAACGAGAA | 1904991–1904972 |
| Z2148 Fwd | Z2148 | Unknown protein encoded within prophage CP-933O | AATATTTTCCCACGTCCCGT | 1927623–1927604 |
| Z2148 Rev | | | GACAGGGTAATCACGCAGGT | 1927516–1927535 |
| Z2149 Fwd | Z2149 | Hypothetical protein | CTTTCGGACTCAGGTGAAGC | 1927995–1928014 |
| Z2149 Rev | | | GGTAAGCCTTCACCAACCAA | 1928224–1928205 |
| Z2150 Fwd | Z2150 | Unknown protein encoded by prophage CP-933R | ACCCCAGTCTCAGGTGTGTC | 2117258–2117277 |
| Z2150 Rev | | | CGGTGAGAGGAGGCTTACTG | 2117540–2117521 |
| Z2366 Fwd | Z2366 | Unknown protein encoded within prophage CP-933R | GAAGAACAGAACCCGCAGAA | 2138525–2138544 |
| Z2366 Rev | | | TGTCTGAACAGGTGCTCTGG | 2138730–2138711 |
| Z2979 Fwd | Z2979 | Putative stability/partitioning protein encoded within prophage CP-933T | CACACCACAGGAAGTGGATG | 2677781–2677800 |
| Z2979 Rev | | | TCCTTTTCACGAATAACCGC | 2678275–2678256 |
| Z3931 Fwd | Z3931 | Unknown protein encoded by prophage CP-933Y | CCGGCACAAACAATAATTCC | 3554499–3554518 |
| Z3931 Rev | | | AGGTGTTGATGATGTTGCCA | 3554671–3554652 |
| Z3933 Fwd | Z3933 | Serine/threonine protein phosphatase | GGCAAGAGCTTTAGCCAGAA | 3556023–3556042 |
| Z3933 Rev | | | AATCACATTCCCCTGGTTCA | 3556182–3556163 |
| Z4315 Fwd | Z4315 | Hypothetical protein | TGCTTCCGTTCTTATTTGCC | 3921831–3921850 |
| Z4315 Rev | | | GCATAAAGGATGGCGTCAAG | 3922024–3922005 |
| Z4325 Fwd | Z4325 | Putative enterotoxin | TATCAGGCCCTTCAAAATGG | 3929618–3929637 |
| Z4325 Rev | | | TATTGCCAAGTACGCCACAA | 3929831–3929812 |
| Z6020 Fwd | Z6020 | Hypothetical protein | CTTCCACGAGGCATTTCATT | 2281021–2281040 |
| Z6020 Rev | | | ACTGATGAGGTTGTCGGAGC | 2281237–228121 |
| Z6021 Fwd | Z6021 | Putative flagellar L-ring protein precursor | AGAAACCTGACTTCGCCTGA | 2282433–2282414 |
| Z6021 Rev | | | CCAGCTTTCCTCCGTGATAA | 2281605–2281624 |
| Z6044 Fwd | Z6044 | Putative major head protein/prohead proteinase | CTTCGCATGACTGAAGGTGA | 2304946–2304965 |
| Z6044 Rev | | | CTGGGAACTGATGGAGGTGT | 2305822–2305803 |

[a] Primer locations and nomenclature of the ORFs refer to nucleotides in the sequence of STEC O157 strain EDL933 (GenBank accession no. AE005174); nt, nucleotide.

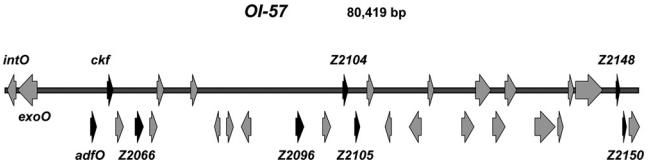



FIG. 1. Map of OI-57. The ORFs investigated in this study (Table 3) are indicated. The *intO* gene indicates the beginning of the genomic island. The island architecture and the nomenclature of the ORFs refer to the STEC O157:H7 strain EDL933 sequence (GenBank accession no. AE005174).

presence of GIs specifically associated with the most pathogenic SPTs, types A and B. The remaining 26 ORFs (Table 2), located in 12 different GIs according to the sequence of the STEC O157 strain EDL933 (25), were considered genes apparently specific for SPTs A and B. To verify this possible association, their presence was further investigated in a larger panel of STEC strains belonging to different SPTs by using specific PCR assays.

**PCR validation of the microarray results.** The panel included STEC strains belonging to SPTs from A to D and E and also *E. coli* strains from healthy human subjects included in the ECOR collection (23). The PCR primers specific for the 26 ORFs apparently associated with SPTs A and B are shown in Table 1. The results of the PCR analyses (Table 2) showed that eight ORFs, located in five different GIs according to the EDL933 strain sequence (25), were present only in strains belonging to SPTs A and B.

One of these ORFs (Z4325) was part of the OI-122 PAI, which has already been shown to be associated with SPTs A and B (9, 19). ORF Z1142 was part of the operon encoding the components of the urea metabolic pathway and related to the *ureD* gene from *Klebsiella aerogenes*. A similar operon is located on a large GI that is present in two copies (OI-43 and OI-48) in the EDL933 strain (24). OI-43 and OI-48 also contain other 4 of the 26 investigated ORFs (Z1154, Z1174, Z1214, and Z1611), which were also present in strains belonging to SPT C and SPTs D and E and in strains from the ECOR collection (Table 2).

ORF Z6020 was specific to SPTs A and B. However, it was located in OI-71 (25), which also contained the ORFs Z6044 and Z6021, frequently observed among strains belonging to SPT C and SPTs D and E, and Z6044, found in 9 of 27 ECOR collection isolates.

The remaining ORFs exclusively present in SPTs A and B were located in OI-57 (Z2097, Z2098, Z2121, and Z2149) and in OI-52 (Z1914). These GIs appeared to be specific to SPT A (OI-52) and SPTs A and B (OI-57), and *in silico* sequence analyses showed that both carry putative virulence genes. In particular, OI-52 harbors the genes conferring resistance to the herbicide methylviologen (Z1870) and encoding the RUS endodeoxyribonuclease (Z1873) (25). OI-57 carries *adfO*, a gene coding for a factor that has been shown to promote the

TABLE 2. PCR detection of the 26 ORFs identified by microarray analysis in STEC strains belonging to seropathotypes (SPTs) A, B, C and D/E and in nonpathogenic *E. coli* strains from the ECOR collection[a]

| Target gene | Genomic island | No. of PCR-positive strains among seropathotype(s)/serogroup(s) (total no. tested): | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SPT A (O157) (9) | SPT B | | | | | SPT C | | SPT D/E (various O groups) (7) | ECOR (27) |
| | | | O26 (9) | O111 (10) | O145 (7) | O103 (4) | O121 (2) | O113 (9) | O91 (4) | | |
| Z0953 | OI-36 | 9 | 8 | 10 | 7 | 4 | 2 | 1 | 1 | 5 | 5 |
| Z0958 | OI-36 | 9 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 5 | 0 |
| Z0989 | OI-36 | 9 | 9 | 10 | 7 | 4 | 0 | 0 | 0 | 5 | 0 |
| Z1142 | OI-43 | 9 | 9 | 10 | 7 | 2 | 2 | 0 | 0 | 0 | 0 |
| Z1154 | OI-43 | 9 | 9 | 10 | 7 | 2 | 2 | 0 | 1 | 0 | 4 |
| Z1174 | OI-43 | 9 | 9 | 10 | 7 | 2 | 2 | 1 | 0 | 1 | 0 |
| Z1214 | OI-43 | 9 | 9 | 10 | 7 | 2 | 0 | 0 | 0 | 6 | 15 |
| Z1348 | OI-44 | 9 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 5 | 0 |
| Z1368 | OI-44 | 9 | 9 | 10 | 7 | 4 | 2 | 2 | 4 | 0 | 9 |
| Z1385 | OI-44 | 9 | 8 | 10 | 7 | 4 | 2 | 0 | 0 | 4 | 0 |
| Z1611 | OI-48 | 9 | 9 | 10 | 7 | 2 | 2 | 1 | 0 | 1 | 0 |
| Z1780 | OI-50 | 9 | 9 | 10 | 7 | 2 | 2 | 1 | 1 | 0 | 2 |
| Z1914 | OI-52 | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Z2097 | OI-57 | 7 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 0 | 0 |
| Z2098 | OI-57 | 7 | 9 | 10 | 7 | 2 | 2 | 0 | 0 | 0 | 0 |
| Z2121 | OI-57 | 9 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 0 | 0 |
| Z2149 | OI-57 | 9 | 8 | 9 | 2 | 2 | 0 | 0 | 0 | 0 | 0 |
| Z2366 | OI-36 | 9 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 5 | 0 |
| Z2979 | OI-76 | 9 | 9 | 10 | 4 | 4 | 2 | 0 | 1 | 4 | 2 |
| Z3931 | OI-108 | 9 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 5 | 0 |
| Z3933 | OI-108 | 9 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 5 | 0 |
| Z4315 | OI-122 | 9 | 6 | 10 | 6 | 2 | 2 | 5 | 0 | 1 | 2 |
| Z4325 | OI-122 | 9 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 0 | 0 |
| Z6021 | OI-71 | 9 | 9 | 10 | 6 | 4 | 2 | 0 | 0 | 3 | 0 |
| Z6020 | OI-71 | 9 | 7 | 2 | 7 | 4 | 2 | 0 | 0 | 0 | 0 |
| Z6044 | OI-71 | 9 | 9 | 10 | 7 | 4 | 2 | 2 | 4 | 0 | 9 |

[a] The number of strains tested is reported in parentheses at the corresponding serogroup heading. The nomenclature of the target genes and genomic islands refers to the sequence of STEC O157 strain EDL933 (GenBank accession no. AE005174).

TABLE 3. PCR analysis of OI-57 in STEC strains belonging to seropathotypes (SPT) A and B[a]

| ORF | Gene product | No. of positive strains (total no. tested) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | SPT A (O157) (9) | SPT B | | | | | SPT C | SPT D/E (various O groups) (7) | ECOR (27) |
| | | | O26 (9) | O111 (10) | O145 (7) | O103 (4) | O121 (2) | O113 (9) | O91 (4) | | |
| Z2053 | AdfO | 9 | 9 | 9 | 7 | 2 | 2 | 0 | 0 | 0 | 0 |
| Z2054 | Ckf | 9 | 9 | 9 | 7 | 2 | 2 | 0 | 0 | 0 | 0 |
| Z2066 | Hypothetical protein | 9 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 0 | 7 |
| Z2096 | Hypothetical protein | 7 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 0 | 0 |
| Z2104 | Putative ARAC-type regulatory protein | 9 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 0 | 2 |
| Z2105 | Hypothetical protein | 7 | 9 | 10 | 7 | 4 | 2 | 2 | 0 | 0 | 2 |
| Z2148 | Hypothetical protein | 9 | 9 | 10 | 7 | 4 | 2 | 0 | 0 | 0 | 0 |
| Z2150 | Hypothetical protein | 9 | 9 | 8 | 1 | 3 | 2 | 0 | 0 | 0 | 0 |

[a] The target genes represented eight ORFs scattered along the entire OI-57 genomic island. STEC strains belonging to SPTs C and D/E and nonpathogenic *E. coli* strains from the ECOR collection were also examined. The number of strains tested is reported in parentheses at the corresponding serogroup. The nomenclature of the ORFs refers to the sequence of STEC O157 strain EDL933 (GenBank accession no. AE005174).

adherence of STEC O157 to *in vitro*-cultured cells (8). OI-57 also carries a gene, hereby termed *ckf*, encoding a putative phage-associated killer protein. Paralogues of this protein have been reported to disrupt the bacterial host membranes when produced in excess (25).

**Analysis of OI-57 in STEC strains.** The apparent association of OI-57 with both SPTs A and B prompted us to investigate its structure in the positive strains. These were analyzed by PCR for the presence of the above-mentioned virulence gene *adfO* as well as for the presence *ckf* and of six additional ORFs scattered along the entire GI and selected for their positions (Table 3). The presence of these eight ORFs was also investigated in STEC strains belonging to SPT C and SPTs D and E and in the ECOR strains, which had been negative for the OI-57 ORFs previously investigated. The results of the PCR assays (Table 3) indicated a strong association of the entire OI-57 with SPTs A and B. In particular, the virulence gene *adfO* was present, always together with *ckf*, in 92.7% of the strains belonging to these SPTs, while both genes were absent from STEC strains belonging to SPT C and SPTs D and E and from the nonpathogenic strains from the ECOR collection. As for the other OI-57 ORFs investigated (Fig. 1), most of them were present in the majority of the SPT A and B strains, suggesting the presence of an entire OI-57. Conversely, these ORFs were rarely observed among the SPT C strains and in ECOR collection isolates (Table 3).

**Presence of OI-57 in other *E. coli* pathogroups.** In order to verify the specificity of the association of OI-57 with STEC SPTs A and B, we investigated its presence in a set of *E. coli* strains belonging to other pathogroups. The *E. coli* isolates included 42 EPEC, 9 ETEC, 8 EAEC, and 8 NTEC strains, which were tested by PCR using primers specific for *adfO* and *ckf* (Table 1). The results of the PCR experiments (Table 4) showed that both genes were present in 71% of EPEC strains, while they could not be detected in the other *E. coli* pathogroups, with the exception of one EAEC strain. In particular, OI-57 was present in three of the eight typical EPEC strains and in 27 of the 34 atypical EPEC strains (Table 5).

**Analysis of OI-57 in EPEC strains.** The structure of OI-57 in the positive EPEC strains was investigated by assessing the presence of the additional six OI-57-associated ORFs (Fig. 1). Variability higher than that observed in STEC was detected, with only 7 of the 30 OI-57-positive EPEC strains harboring the complete set of ORFs. Among the remaining strains, 10 were positive for five of the ORFs (9 for Z2066, Z2104, Z2105, Z2148, and Z2150 and 1 with ORF Z2096 instead of ORF Z2105), 7 were positive for four ORFs (4 strains possessed ORFs Z2066, Z2104, Z2148, and Z2150 and 3 strains had ORF Z2105 instead of ORF Z2150), 4 possessed ORFs Z2066, Z2104, and Z2148, and 2 strains were positive for ORF Z2148 only.

## DISCUSSION

The pathogenicity of STEC represents a complex phenomenon that has not yet been fully elucidated. Human beings are exposed to a wide variety of STEC strains, which can be frequently found in animal reservoirs, in foodstuffs of animal origin, and in the environment. However, most of these strains, albeit able to produce Stx, are rarely associated with severe human disease (3), probably because they lack other virulence factors needed for full pathogenicity. Therefore, the strains pathogenic to humans appear to be a subpopulation of STEC and have been categorized into four seropathotypes (from A to D), according to serotype, relative frequencies in human disease, association with outbreaks and HUS, and other genomic features (9, 11, 12). An additional seropathotype (SPT E)

TABLE 4. PCR detection of OI-57 *adfO* and *ckf* genes in *E. coli* strains belonging to different pathogroups[a]

| *E. coli* pathogroup (no. of strains) | No. of strains with the PCR profile | PCR amplification of[b]: | |
|---|---|---|---|
| | | ORF Z2053[c] | ORF Z2054[d] |
| EPEC (42) | 30 | + | + |
| | 12 | − | − |
| EAEC (8) | 7 | − | − |
| | 1 | + | + |
| ETEC (9) | 9 | − | − |
| NTEC (8) | 8 | − | − |
| ECOR (27) | 27 | − | − |

[a] The nomenclature of the ORFs refers to the sequence of STEC O157 strain EDL933 (GenBank accession no. AE005174).
[b] +, amplified; −, not amplified.
[c] *adfO*, encoding the intestinal colonization factor.
[d] *ckf*, encoding the cell killing factor.

TABLE 5. Presence or absence of OI-57 in typical (EAF plasmid-positive) and atypical (EAF plasmid-negative) EPEC by serogroup

| Serogroup | No. of strains with the PCR profile/total no. of strains tested | PCR amplification of[a]: | |
|---|---|---|---|
| | | OI-57[b] | EAF |
| O125 | 2/3 | + | − |
| O125 | 1/3 | − | − |
| O26 | 12/12 | + | − |
| O45 | 1/1 | + | − |
| O55 | 3/4 | + | − |
| O55 | 1/4 | − | + |
| O86 | 2/2 | − | + |
| O103 | 2/2 | + | − |
| O111 | 2/4 | + | + |
| O111 | 1/4 | + | − |
| O111 | 1/4 | − | − |
| O114 | 1/1 | + | + |
| O127 | 2/5 | − | + |
| O127 | 3/5 | − | − |
| O128 | 4/6 | + | − |
| O128 | 2/6 | − | − |
| O ND[c] | 2/2 | + | − |

[a] +, amplified; −, not amplified.
[b] Positivity with respect to OI-57 implies the presence of both the *adfO* and *ckf* genes.
[c] O ND, not determined.

includes STEC strains that seem to be confined to animal hosts (9, 11).

The introduction of the pathogenomic approach, based on the use of whole-genome high-throughput technologies, led to a deeper knowledge of the virulence characteristics of STEC strains and showed that their full pathogenicity may result from the activation of complex pathways involving the action of several virulence factors encoded by MGEs. These include the PAIs LEE (16) and OI-122 (9, 19) and the large virulence plasmids carrying the operon encoding the enterohemolysin (28). These MGEs are consistently present in STEC strains belonging to the serogroups most frequently associated with HUS and included in SPTs A and B and can undoubtedly be considered part of the genomic asset that composes the STEC "virulome." Other factors that have been believed to play a role in STEC virulence and pathogenicity include fimbriae (29) and effectors translocated via the LEE-encoded type 3 secretion system (15, 16) or the plasmid-borne type 2 secretion system (8). However, although the data provided were in most cases conclusive, the presence of these factors has not been investigated in a large number of strains belonging to the different SPTs.

In recent years, genome sequencing projects have provided an important contribution to the characterization of the genomic arrays needed for the virulence of bacterial pathogens. Sequencing of the STEC O157 chromosome showed the presence of 177 GIs that are absent in *E. coli K-12* and that, in many cases, possess the characteristics of putative PAIs (7, 25), even though their actual roles in virulence are often still unknown.

To define the genomic map of the traits characterizing the pathogenic STEC SPTs, we used microarray hybridization to compare the whole genome of STEC O157 with the genomes of a panel of STEC strains encompassing SPTs B, C, and D. The comparison identified some ORFs, located on GIs, that appeared to be associated with SPTs A and B; these were further investigated using PCR amplification and a larger panel of *E. coli* strains.

This approach led us to identify a GI termed OI-57 in the EDL933 strain chromosome (25) that was consistently present in STEC strains belonging to the highly pathogenic SPTs A and B but not in the less pathogenic strains of SPT C and SPTs D and E. OI-57 harbors two adjacent genes whose products are annotated with a putative function: *adfO*, encoding a putative virulence factor promoting the adherence of STEC O157 to HeLa cells (8), and *ckf*, encoding a phage-associated putative bacterial cell-killing factor (25). The latter has a paralogous protein that acts by disrupting the bacterial host membranes when produced in excess, as reported in GenBank (accession no. AE005174).

The analysis of *E. coli* isolates belonging to other pathogroups showed that OI-57 was also present in a high proportion of human EPEC strains belonging to the most common serogroups. EPEC strains can be classified into two subtypes (typical or atypical) according to the presence or absence of bundle-forming pili (35), which are fimbrial adhesins that represent a marker of the EAF plasmid (6). We could observe a higher rate of the presence of OI-57 in atypical EPEC (Table 5). However, due to the conflicting reports on the pathogenicity of atypical EPEC (35) and to the low number of typical EPEC strains examined, associations of OI-57 with the EPEC clones cannot be inferred.

The presence of OI-57 in EPEC strains is still interesting, since they share with STEC strains belonging to SPTs A and B an important pathogenetic mechanism: the capability to colonize the host intestinal mucosa, inducing the peculiar lesion known as "attaching and effacing" (20). This mechanism is governed by the presence of the LEE (16, 20), and the strong association of OI-57 with LEE-positive strains suggests its possible involvement in the gut colonization processes. This hypothesis is supported by experimental observations showing that the *adfO* gene enhanced the adhesivity of STEC O157 to *in vitro*-cultured cells (8) and was upregulated in STEC O157 lineages characterized by a high attack rate in outbreaks (1).

Despite the strong association of OI-57 with the most pathogenic STEC seropathotypes, we could not observe any association with a particular lineage of LEE-positive *E. coli* (35). This finding may reflect a general synergism between the products of the genes harbored by these two mobile genetic elements in the colonization of the host gut.

According to the hypotheses of dynamic modeling of the genomes of the pathogens (5, 14), PAIs usually assemble genes derived from plasmids and phages in cointegrated structures that are then stabilized by selective pressure through the inactivation of the mobility and recombination genes or are removed by loss of genetic material by recombination events (13). To evaluate whether OI-57 can be considered a stable acquisition of the STEC genome, we investigated its integrity by PCR amplification of eight ORFs, encompassing the length of the whole island. Our PCR analyses showed that a presumably entire OI-57 was present in the majority of the positive STEC strains examined (Table 3), thus indicating that the island represents a stable acquisition in these clonal lineages.

The analysis of OI-57 integrity in EPEC, the only other *E. coli* pathogroup that appears to have acquired the island, showed a higher degree of variability than in STEC, with less than one-fourth of the strains showing a complete collection of the ORFs tested.

In conclusion, our results indicate that the STEC strains causing severe disease in humans and belonging to SPTs A and B possess OI-57 as part of their virulome. OI-57 could be involved in the attaching-and-effacing colonization of the intestinal mucosa and appears to be present in its entirety in the majority of the STEC strains investigated.

Further studies are needed to elucidate the roles of the putative virulence gene *adfO* and of *ckf* in the pathogenesis of STEC infections. These genes could also represent candidate targets for the development of diagnostic tools aimed at the identification of the STEC strains constituting a threat to human health among the wealth of Stx-producing strains present in animal reservoirs and in vehicles of infections.

## ACKNOWLEDGMENTS

This work was supported by European Union grant SANCO 2005/FOOD SAFETY/031 Public Health-*E. coli* from the EU Commission, by grant RFPS-2007-7-638803.P1 from the Italian Ministry of Health, and by grant FPI 20060054361 from Spanish Ministry of Education and Science.

We are solely responsible for the work described in this paper, and our opinions are not necessarily those of the E.U.

## REFERENCES

1. **Abu-Ali, G. S., L. M. Ouellette, S. T. Henderson, T. S. Whittam, and S. D. Manning.** 2010. Differences in adherence and virulence gene expression between two outbreak strains of enterohaemorrhagic *Escherichia coli* O157:H7. Microbiology **156:**408–419.
2. **Brooks, J. T., E. G. Sowers, J. G. Wells, K. D. Greene, P. M. Griffin, R. M. Hoekstra, and N. A. Strockbine.** 2005. Non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States, 1983–2002. J. Infect. Dis. **192:**1422–1429.
3. **Caprioli, A., S. Morabito, H. Brugère, and E. Oswald.** 2005. Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission. Vet. Res. **36:**289–311.
4. **Caprioli, A., A. E. Tozzi, G. Rizzoni, and H. Karch.** 1997. Non-O157 Shiga toxin-producing *Escherichia coli* infections in Europe. Emerg. Infect. Dis. **3:**578–579.
5. **de la Cruz, F., and J. Davies.** 2000. Horizontal gene transfer and the origin of species: lessons from bacteria. Trends Microbiol. **8:**128–133.
6. **Franke, J., S. Franke, H. Schmidt, A. Schwarzkopf, L. H. Wieler, G. Baljer, L. Beutin, and H. Karch.** 1994. Nucleotide sequence analysis of enteropathogenic Escherichia coli (EPEC) adherence factor probe and development of PCR for rapid detection of EPEC harbouring virulence plasmids. J. Clin. Microbiol. **32:**2460–2463.
7. **Hayashi, T., K. Makino, M. Ohnishi, K. Kurokawa, K. Ishii, K. Yokoyama, C. G. Han, E. Ohtsubo, K. Nakayama, T. Murata, M. Tanaka, T. Tobe, T. Iida, H. Takami, T. Honda, C. Sasakawa, N. Ogasawara, T. Yasunaga, S. Kuhara, T. Shiba, M. Hattori, and H. Shinagawa.** 2001. Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. DNA Res. **8:**11–22.
8. **Ho, T. D., B. M. Davis, J. M. Ritchie, and M. K. Waldor.** 2008. Type 2 secretion promotes enterohemorrhagic *Escherichia coli* adherence and intestinal colonization. Infect. Immun. **76:**1858–1865.
9. **Karmali, M. A., M. Mascarenhas, S. Shen, K. Ziebell, S. Johnson, R. Reid-Smith, J. Isaac-Renton, C. Clark, K. Rahn, and J. B. Kaper.** 2003. Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. J. Clin. Microbiol. **41:**4930–4940.
10. **Klapproth, J. M., I. C. Scaletsky, B. P. McNamara, L. C. Lai, C. Malstrom, S. P. James, and M. S. Donnenberg.** 2000. A large toxin from pathogenic *Escherichia coli* strains that inhibits lymphocyte activation. Infect. Immun. **68:**2148–2155.
11. **Konczy, P., K. Ziebell, M. Mascarenhas, A. Choi, C. Michaud, A. M. Kropinski, T. S. Whittam, M. Wickham, B. Finlay, and M. A. Karmali.** 2008. Genomic O island 122, locus for enterocyte effacement, and the evolution of virulent verocytotoxin-producing *Escherichia coli*. J. Bacteriol. **190:**5832–5840.
12. **Kresse, A. U., I. Rienacker, A. M. Valle, H. Steinruck, H. Claus, S. M. Payne, H. Tschape, P. H. Williams, and R. Reissbrodt.** 2007. Enterohaemorrhagic *Escherichia coli* O157 and non-O157 serovars differ in their mechanisms for iron supply. Int. J. Med. Microbiol. **297:**9–15.
13. **Lawrence, J. G., and H. Ochman.** 1998. Molecular archaeology of the *Escherichia coli* genome. Proc. Natl. Acad. Sci. U. S. A. **95:**9413–9417.
14. **Lederberg, J.** 1997. Infectious disease as an evolutionary paradigm. Emerg. Infect. Dis. **3:**417–423.
15. **Loukiadis, E., R. Nobe, S. Herold, C. Tramuta, Y. Ogura, T. Ooka, S. Morabito, M. Kerouredan, H. Brugere, H. Schmidt, T. Hayashi, and E. Oswald.** 2008. Distribution, functional expression, and genetic organization of Cif, a phage-encoded type III-secreted effector from enteropathogenic and enterohemorrhagic *Escherichia coli*. J. Bacteriol. **190:**275–285.
16. **McDaniel, T. K., and J. B. Kaper.** 1997. A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. Mol. Microbiol. **23:**399–407.
17. **Morabito, S., H. Karch, H. Schmidt, F. Minelli, P. Mariani-Kurkdjian, F. Allerberger, K. A. Bettelheim, and A. Caprioli.** 1999. Molecular characterisation of verocytotoxin-producing *Escherichia coli* of serogroup O111 from different countries. J. Med. Microbiol. **48:**891–896.
18. **Morabito, S., R. Tozzoli, A. Caprioli, H. Karch, and A. Carattoli.** 2002. Detection and characterization of class 1 integrons in enterohemorrhagic *Escherichia coli*. Microb. Drug Resist. **8:**85–91.
19. **Morabito, S., R. Tozzoli, E. Oswald, and A. Caprioli.** 2003. A mosaic pathogenicity island made up of the locus of enterocyte effacement and a pathogenicity island of *Escherichia coli* O157:H7 is frequently present in attaching and effacing *E. coli*. Infect. Immun. **71:**3343–3348.
20. **Nataro, J. P., and J. B. Kaper.** 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. **11:**142–201.
21. **Nicholls, L., T. H. Grant, and R. M. Robins-Browne.** 2000. Identification of a novel genetic locus that is required for *in vitro* adhesion of a clinical isolate of enterohaemorrhagic *Escherichia coli* to epithelial cells. Mol. Microbiol. **35:**275–288.
22. **O'Brien, A. D., J. W. Newland, S. F. Miller, R. K. Holmes, H. W. Smith, and S. B. Formal.** 1984. Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. Science **226:**694–696.
23. **Ochman, H., and R. K. Selander.** 1984. Standard reference strains of *Escherichia coli* from natural populations. J. Bacteriol. **157:**690–693.
24. **Oswald, E., H. Schmidt, S. Morabito, H. Karch, O. Marches, and A. Caprioli.** 2000. Typing of intimin genes in human and animal enterohemorrhagic and enteropathogenic *Escherichia coli*: characterization of a new intimin variant. Infect. Immun. **68:**64–71.
25. **Perna, N. T., G. Plunkett III, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner.** 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature **409:**529–533.
26. **Robins-Browne, R. M., and E. L. Hartland.** 2002. *Escherichia coli* as a cause of diarrhea. J. Gastroenterol. Hepatol. **17:**467–475.
27. **Saeed, A. I., V. Sharov, J. White, J. Li, W. Liang, N. Bhagabati, J. Braisted, M. Klapa, T. Currier, M. Thiagarajan, A. Sturn, M. Snuffin, A. Rezantsev, D. Popov, A. Ryltsov, E. Kostukovich, I. Borisovsky, Z. Liu, A. Vinsavich, V. Trush, and J. Quackenbush.** 2003. TM4: a free, open-source system for microarray data management and analysis. Biotechniques **34:**374–378.
28. **Schmidt, H., L. Beutin, and H. Karch.** 1995. Molecular analysis of the plasmid-encoded hemolysin of *Escherichia coli* O157:H7 strain EDL 933. Infect. Immun. **63:**1055–1061.
29. **Shen, S., M. Mascarenhas, R. Morgan, K. Rahn, and M. A. Karmali.** 2005. Identification of four fimbria-encoding genomic islands that are highly specific for verocytotoxin-producing *Escherichia coli* serotype O157 strains. J. Clin. Microbiol. **43:**3840–3850.
30. **Stevens, M. P., A. J. Roe, I. Vlisidou, P. M. van Diemen, R. M. La Ragione, A. Best, M. J. Woodward, D. L. Gally, and T. S. Wallis.** 2004. Mutation of *toxB* and a truncated version of the *efa-1* gene in *Escherichia coli* O157:H7 influences the expression and secretion of locus of enterocyte effacement-encoded proteins but not intestinal colonization in calves or sheep. Infect. Immun. **72:**5402–5411.
31. **Stevens, M. P., P. M. van Diemen, G. Frankel, A. D. Phillips, and T. S. Wallis.** 2002. Efa1 influences colonization of the bovine intestine by Shiga toxin-producing *Escherichia coli* serotypes O5 and O111. Infect. Immun. **70:**5158–5166.
32. **Tarr, P. I., C. A. Gordon, and W. L. Chandler.** 2005. Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. Lancet **365:**1073–1086.
33. **Tarr, C. L., T. M. Large, C. L. Moeller, D. W. Lacher, P. I. Tarr, D. W.**

**Acheson, and T. S. Whittam.** 2002. Molecular characterization of a serotype O121:H19 clone, a distinct Shiga toxin-producing clone of pathogenic *Escherichia coli*. Infect. Immun. **70:**6853–6859.
34. **Tatsuno, I., M. Horie, H. Abe, T. Miki, K. Makino, H. Shinagawa, H. Taguchi, S. Kamiya, T. Hayashi, and C. Sasakawa.** 2001. *toxB* gene on pO157 of enterohemorrhagic *Escherichia coli* O157:H7 is required for full epithelial cell adherence phenotype. Infect. Immun. **69:**6660–6669.
35. **Tennant, S. M., M. Tauschek, K. Azzopardi, A. Bigham, V. Bennett-Wood, E. L. Hartland, W. Qi, T. S. Whittam, and R. M. Robins-Browne.** 2009. Characterisation of atypical enteropathogenic *E. coli* strains of clinical origin. BMC Microbiol. **9:**117.
36. **Tozzi, A. E., A. Caprioli, F. Minelli, A. Gianviti, L. De Petris, A. Edefonti, G. Montini, A. Ferretti, T. De Palo, M. Gaido, and G. Rizzoni.** 2003. Shiga toxin-producing *Escherichia coli* infections associated with hemolytic uremic syndrome, Italy, 1988–2000. Emerg. Infect. Dis. **9:**106–108.
37. **Tozzoli, R., A. Caprioli, S. Cappannella, V. Michelacci, M. L. Marziano, and S. Morabito.** 2010. Production of the subtilase AB5 cytotoxin by Shiga toxin-negative *Escherichia coli*. J. Clin. Microbiol. **48:**178–183.
38. **Tozzoli, R., A. Caprioli, and S. Morabito.** 2005. Detection of *toxB*, a plasmid virulence gene of *Escherichia coli* O157, in enterohemorrhagic and enteropathogenic *E. coli*. J. Clin. Microbiol. **43:**4052–4056.
39. **World Health Organization.** 1999. Zoonotic non-O157 Shiga toxin-producing *Escherichia coli* (STEC). Report of a WHO scientific working group meeting. World Health Organization, Geneva, Switzerland.

*Editor:* B. A. McCormick

# 5. Analysis of VTEC strains possessing rare combinations of MGEs vehiculating virulence factors

# 5. Analysis of VTEC strains possessing rare combinations of MGEs vehiculating virulence factors

*Introduction*

In 2011, Germany experienced the largest outbreak of VTEC infections ever recorded: a total of 3,842 cases were reported, including 855 cases of haemolytic uremic syndrome (HUS) and a total of 53 fatal outcomes. The outbreak involved an unusually high proportion of case-patients with haemolytic uremic syndrome, with many adults affected and frequent development of neurologic symptoms. The strain responsible for the outbreak was an O104:H4 Enteroaggregative *E. coli* strain which had acquired the ability to produce verocytotoxins. Due to these peculiar characteristics, no methods were available for the specific detection of such a strain when the outbreak occurred.

A second point-source outbreak caused by the same VT-producing EAEC O104 strain was reported in June 2011 in France. Both the outbreaks were linked to eating fenugreek sprouts obtained from seeds produced in Egypt and distributed in Germany and other European countries.

*Results*

The prospective reviewing of the *E. coli* strains culture collection held at the Istituto Superiore di Sanità allowed the identifification of an O104:H4 VT-producing EAEC strain, termed ED 703, which had been isolated in Italy in 2009 from a child affected by bloody diarrhoea, followed by HUS and neurological symptoms. Pulsed-field-gel electrophoresis (PFGE) showed a 94.7% of similarity with the strain responsible for the German outbreak. An epidemiological interview was carried on with the family of the child and the acquisition of the infection was traced back to a recent travel in Tunisia probably through an intra-family person to person transmission. (Publication n.1: *Emerging Infectious Diseases*, 17(10):1957-1958. Impact Factor 6.169).

In order to provide advice to the laboratories involved in diagnostic analysis and food testing during the epidemiological surveys, a molecular method has been developed for the detection of O104:H4 VTEC strains in the test matrices considered as suspected vehicles of infection. Real Time-PCR tools targeting the O104 and H4 antigens-associated genes (*wzxO104* and *fliCH4*) have been developed to be used in association with the *vtx*-specific reagents previously available as part of the standard method ISO/TS 13136. This method allows the direct rapid determination of the absence of O104:H4 VTEC strains in food matrices in up to 24 hours and the identification of a presumptive contamination by these strains in up to 29 hours. (Publication n.2: http://www.iss.it/binary/vtec/cont/Lab_proc_O104_rev2.pdf).

*Conclusions*

The results obtained in this study support the hypothesis that ruminants would not have had a specific role in the transmission of O104:H4 VT-producing EAEC strain during the German outbreak and confirm that the unusual combination of virulence factors of VTEC and EAEC pathotypes can confer a high degree of virulence to these strains. Furthermore, the high similarity assessed between the two strains together with

the acquisition of the infections from underdeveloped countries of the globe suggest the stability of this pathogenic clone in such geographical areas.

The proposed method represents an important tool for the fast and specific detection of this rare and unusual VTEC serotype in foodstuff and during the German outbreak it served as an effective rapid method for the epidemiological surveys conducted with the aim of tracing back the vehicle of the infections by the outbreak strains.

# Similarity of Shiga Toxin–producing *Escherichia coli* O104:H4 Strains from Italy and Germany

**To the Editor:** Since the beginning of May 2011, a large outbreak of infections associated with Shiga toxin (Stx)–producing *Escherichia coli* (STEC) O104:H4 has occurred in Germany (*1*). The outbreak showed 3 unusual features: 1) a large proportion of case-patients with hemolytic uremic syndrome (HUS); 2) HUS in adults, although it usually affects children; and 3) frequent development of neurologic symptoms in patients when clinical and laboratory markers of HUS were improving (*1*,*2*). A second point-source outbreak caused by the same STEC O104 strain was reported in June 2011 in France (*3*). Both outbreaks were linked to eating fenugreek sprouts obtained from seeds produced in Egypt and distributed in Germany and other European countries (*4*).

Instead of the attaching–effacing mechanism of adhesion to intestinal mucosa that is typical of STEC associated with severe human disease (*5*), the STEC O104 epidemic strain had genetic markers and an adhesion pattern (*6*) typical of enteroaggregative *E. coli* (EAEC), another group of diarrheagenic strains found frequently in developing countries (*5*).

On basis of these findings, we reviewed our culture collection and found that an STEC strain (ED-703) from a case-patient with HUS in 2009 in Italy had the same combination of virulence factors as the strain from Germany: Stx2 production and enteroaggregative adhesion genetic markers. This strain, which had not been typed when it was isolated, showed positive PCR results for O104 (*7*) and H4 (*8*) antigen–associated genes and was agglutinated by an O104 antiserum (Statens Serum Institut, Copenhagen, Denmark). Pulsed-field gel electrophoresis showed a high degree of similarity (94.7%) with the outbreak strain from Germany (provided by M. Mielke, Robert Koch Institute, Berlin, Germany). In contrast with the outbreak strain, ED-703 did not produce extended-spectrum β-lactamases.

The strain from our culture collection had been isolated from a 9-year-old girl admitted to the pediatric nephrology unit of the Ospedale Maggiore (Milan, Italy) on August 5, 2009, after 5 days of bloody diarrhea, vomiting, and abdominal pain. Diagnosis of HUS was based on the presence of hemolytic anemia, thrombocytopenia, and anuria. Neurologic symptoms (e.g., lethargy, diplopia, and nystagmus) occurred during hospitalization; magnetic resonance imaging showed signal abnormalities in the lenticular nuclei.

Because of severe cardiac impairment with ejection fraction reduction and troponin increase, inotropic support and mechanical ventilation were temporarily needed. After improvement of clinical conditions, the patient was discharged, but she was readmitted a few days later because of headache, vomiting, confusion, dysarthria, hypertension, and visual impairment. Ischemic lesions were found by magnetic resonance imaging at fundus oculi. Neurologic status improved the next day, but the visual deficit persisted. Hemodialysis was needed for 2 months. Long-term sequelae of the disease were stage IV chronic kidney disease, hypertension, and severe visual impairment.

Informed consent and an epidemiologic interview were obtained from the patient's parents. The household, including her mother and 2 siblings (4 and 5 years of age), had traveled for 1 week to a resort in Tunisia; they had returned 3 weeks before the onset of the prodromal symptoms of HUS. Four days after their return, the youngest sister was hospitalized for 3 days because of bloody diarrhea, but no laboratory diagnosis was established. The mother reported having had watery diarrhea and abdominal pain on August 2. The patient history did not show any other usual risk factor for STEC infection, such as consumption of unpasteurized milk or dairy products, undercooked meat, or raw sprouts or direct exposure to ruminants or their manure. This finding suggests that the infection was probably acquired through person-to-person transmission.

This case report confirms that strains of STEC O104 strictly related to the epidemic strain in Germany had already caused sporadic infections in Europe (*9*). Other cases have been documented in 2001 in Germany (*6*,*9*), in 2004 in France (*9*), and in 2010 in Finland in a patient with diarrhea who had traveled to Egypt (*9*). Both of the cases for which the information on the origin of the infection was available were related to travel to northern Africa, from which the seeds associated with both outbreaks could be traced (*4*).

The history of this patient supports the hypothesis that ruminants would not have had a specific role in the transmission of STEC O104:H4, as already suggested by the epidemiologic features of the recent outbreaks (*1*,*3*). In fact, STEC O104 cannot be considered true STEC but rather EAEC strains that acquired the Stx2-coding phages by horizontal gene transfer, and EAEC is considered to be a human pathogen usually transmitted by the oral–fecal route (*5*).

The clinical course of our patient closely resembles those of persons who had HUS associated with the German outbreak (*1*,*2*). The unusual combination of virulence factors of STEC and EAEC, already described in a group of STEC O111:H2 from an outbreak of HUS in France in 1996

(*10*), might confer a high degree of virulence to these strains. It also might explain the severity of the clinical findings associated with STEC O104:H4 infections.

**Gaia Scavia, Stefano Morabito, Rosangela Tozzoli, Valeria Michelacci, Maria Luisa Marziano, Fabio Minelli, Clarissa Ferreri, Fabio Paglialonga, Alberto Edefonti, and Alfredo Caprioli**

Author affiliations: Istituto Superiore di Sanità, Rome, Italy (G. Scavia, S. Morabito, R. Tozzoli, V. Michelacci, M.L. Marziano, F. Minelli, C. Ferreri, A. Caprioli); and Ospedale Maggiore Policlinico, Milan, Italy (F. Paglialonga, A. Edefonti)

DOI: http://dx.doi.org/10.3201/eid1710.111072

### References

1. Frank C, Werber D, Cramer JP, Askar M, Faber M, Heiden MA, et al. Epidemic profile of Shiga toxin–producing *Escherichia coli* O104:H4 outbreak in Germany—preliminary report. N Engl J Med. 2011 June 22; [Epub ahead of print].
2. Jansen A, Kielstein J. The new face of enterohaemorrhagic *Escherichia coli* infections. Euro Surveill. 2011;16:pii:19898.
3. Gault G, Weill FX, Mariani-Kurkdjian P, Jourdan-da Silva N, King L, Aldabe B, et al. Outbreak of haemolytic uraemic syndrome and bloody diarrhoea due to *Escherichia coli* O104:H4, southwest France, June 2011. Euro Surveill. 2011;16:pii:19905.
4. European Food Safety Authority. Tracing seeds, in particular fenugreek (*Trigonella foenum-graecum*) seeds, in relation to the Shiga toxin–producing *E. coli* (STEC) O104:H4 2011 outbreaks in Germany and France [cited 5 Jul 2011]. http://www.efsa.europa.eu/en/supporting/doc/176e.pdf
5. Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli.* Nat Rev Microbiol. 2004;2:123–40. doi:10.1038/nrmicro818
6. Bielaszewska M, Mellmann A, Zhang W, Köck R, Fruth A, Bauwens A, et al. Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. Lancet Infect Dis. 2011 June 22; [Epub ahead of print].
7. Bugarel M, Beutin L, Martin A, Gill A, Fach P. Micro-array for the identification of Shiga toxin–producing *Escherichia coli* (STEC) seropathotypes associated with hemorrhagic colitis and hemolytic uremic syndrome in humans. Int J Food Microbiol. 2010;142:318–29. doi:10.1016/j.ijfoodmicro.2010.07.010
8. EU Reference Laboratory for. *E. coli*. Detection and identification of Verocytotoxin-producing *Escherichia coli* (VTEC) O104:H4 in food by real time PCR—laboratory procedure [cited 2011 Jul 5]. http://www.iss.it/binary/vtec/cont/Lab_proc_O104_rev2.pdf
9. European Centre for Disease Prevention and Control. Shiga toxin/verotoxin–producing *Escherichia coli* in humans, food and animals in the EU/EEA, with special reference to the German outbreak strain STEC O104 [cited 2011 Jul 5]. http://www.ecdc.europa.eu/en/publications/Publications/1106_TER_EColi_joint_EFSA.pdf
10. Morabito S, Karch H, Mariani-Kurkdjian P, Schmidt H, Minelli F, Bingen E, et al. Enteroaggregative, Shiga toxin–producing *Escherichia coli* O111: H2 associated with an outbreak of hemolytic-uremic syndrome. J Clin Microbiol. 1998;36:840–2.

Address for correspondence: Alfredo Caprioli, Istituto Superiore di Sanità, Viale Regina Elena 299, Rome 00161, Italy; email: alfredo.caprioli@iss.it

## Complicated Pandemic (H1N1) 2009 during Pregnancy, Taiwan

**To the Editor:** Pregnant women with pandemic (H1N1) 2009 virus infection are at increased risk for severe illness and complications (*1*–*3*). Recent reports have shown that this infection causes disproportionate illness and death in pregnant women and has been associated with adverse fetal and neonatal outcomes. We characterized the severity of pandemic (H1N1) 2009 virus infection among pregnant women in Taiwan.

Complicated influenza infection, defined as influenza-like illness and evidence of pneumonia, neurologic symptoms, myopericarditis, or invasive bacterial infections, has been a notifiable disease in Taiwan since 2002 (*4*). We reviewed reports and medical records of complicated pandemic (H1N1) 2009 virus infection, confirmed by real-time reverse transcription PCR in women 15–49 years of age who had onset of illness during July 1–December 31, 2009. Data were obtained for demographics; pregnancy status and outcome; gestational age at illness onset; preexisting medical conditions; onset of illness; treatment; and severity, including intensive care unit (ICU) admission.

To calculate rates of complicated pandemic (H1N1) 2009 virus infection, we estimated the pregnant population during July 1–December 31, 2009, by using the National Health Insurance computerized database for Taiwan (*5*). Women who were 15–49 years of age and had been assigned International Classification of Diseases, 9th Revision, Clinical Modification (www.cdc.gov/nchs/icd/icd9cm.htm), codes of V22* (normal pregnancy) and V23* (supervision of high-risk pregnancy) during the study were considered pregnant. Number of nonpregnant women was estimated by subtracting the calculated number of pregnant women from the number of women 15–49 years of age from 2009 household registration data (*6*). We estimated 95% confidence intervals (CIs) for rates by using exact binomial methods.

During July 1–December 31, 2009, data were reported for 10 pregnant women and 138 nonpregnant women 15–49 years of age who had confirmed, complicated pandemic (H1N1) 2009 virus infections. Dates of illness onset ranged from August 3 through December 31, 2009. Median age of the 10 pregnant women was 24.5 years (range 22–32 years), and median

# Detection and identification of Verocytotoxin-producing *Escherichia coli* (VTEC) O104:H4 in food by Real Time PCR

## Laboratory procedure

### Aims and field of application:

The ongoing large outbreak of VTEC infections in Germany is caused by a VTEC strain belonging to serotype O104:H4, a serotype (and serogroup), not comprised among those usually associated with severe infections in Europe and worldwide.

Some characteristics of the outbreak strain have been reported:

1. It produces VT2, and harbors the *vtx2a* gene subtype.
2. It lacks the gene coding for the adherence factor intimin (*eae* gene), which is considered as a hallmark of the pathogenic VTEC.
3. It is multi-resistant: Ampicillin, Cefoxitin, Cefotaxim, Cetfazidim, Streptomycin, Tetracyclin, Trimethoprim/Sulfamethoxazol, Nalidixic acid. It produces an expanded spectrum beta lactamase (ESBL): CTX-M 15.
4. All but one strain tested so far posses the genetic markers of typical of Enteroaggregative *Escherichia coli* (EAggEC): the *aggR, aatA*, *aaiC* and *aap* genes.

These features have to be considered when defining a diagnostic strategy for the detection in food or environmental samples. The absence of the *eae* gene may pose problems, since the ISO TS 13136 - Horizontal method for the detection of Shiga toxin-producing *Escherichia coli* (STEC) belonging to O157, O111, O26, O103 and O145 serogroups - Qualitative Real-time polymerase chain reaction (PCR)-based Method – is based on the stepwise detection of the *eae* gene in *vtx*-positive samples. Serotyping is cumbersome and only few laboratories may possess antiserum specific to this particular serogroup.

Therefore, two genetic markers can be considered for the molecular screening of the *vtx*-positive enrichment cultures: the O104 antigen-associated gene $wzx_{O104}$ and the gene encoding the H4 flagellar antigen, $fliC_{H4}$. The markers associated with the enteroaggregative adhesion could also be considered as targets of the diagnostic procedure and such a test is currently under evaluation at the EU RL VTEC.

The antibiotic resistance characteristics of the VTEC O104 outbreak strain can be exploited for the isolation step, by plating PCR-positive enrichment cultures samples onto MacConkey agar supplemented with streptomycin (20 μg/ml) and/or tetracycline (10 mg/ml), or onto *Brilliance™ ESBL* Agar from Oxoid or ChromID ESBL from Biomerieux or similar media available in the commerce.

The proposed method aims at the identification of the presence of O104 antigen-associated gene ($wzx_{O104}$) in *vtx*-positive enrichment cultures. The molecular design of this Real Time PCR has been described in the literature (Bugarel et al., 2010). The same assay can be used to identify and confirm the O104:H4 serotype of isolated strains in combination with the $fliC_{H4}$ RT-PCR that has been deployed at the EU-RL for *E. coli* on the *fliC* gene sequence of the *E. coli* strain U9-41 present in GenBank under the accession number AY249989.

**Food samples screening - Procedure**

Enrichment cultures are performed by adding a 25 gr test portion of food sample or 25 ml of milk sample to 225 ml of Buffered Peptone Water, and incubating for 18 -24 h at 37°C ± 1°C. One ml aliquot of such a culture is used for DNA extraction and purification. For testing of seeds used for production of sprouts to be consumed raw, refer to **Annex 4**.

This Real Time PCR protocol is used to test all the samples that give positive results for the presence of *vtx* genes by using the first step of the ISO TS

13136 – Horizontal method for the detection of Shiga toxin-producing *Escherichia coli* (STEC) belonging to O157, O111, O26, O103 and O145 serogroups - Qualitative Real-time polymerase chain reaction (PCR)-based Method.

The Real-time PCRs for $wzx_{O104}$ and $fliC_{H4}$ are performed using the primers and probes described in **Annex 1**. Amplification conditions to be applied will depend on the system used, and will refer to the instructions supplied with the instrument and the reagents' kit of choice.

A standard two-step thermal profile has been used at EU RL VTEC with satisfactory results for both the reactions. The details were as follows:

95°C X 10'

35 cycles of

95°C X 15''

60°C X 1'

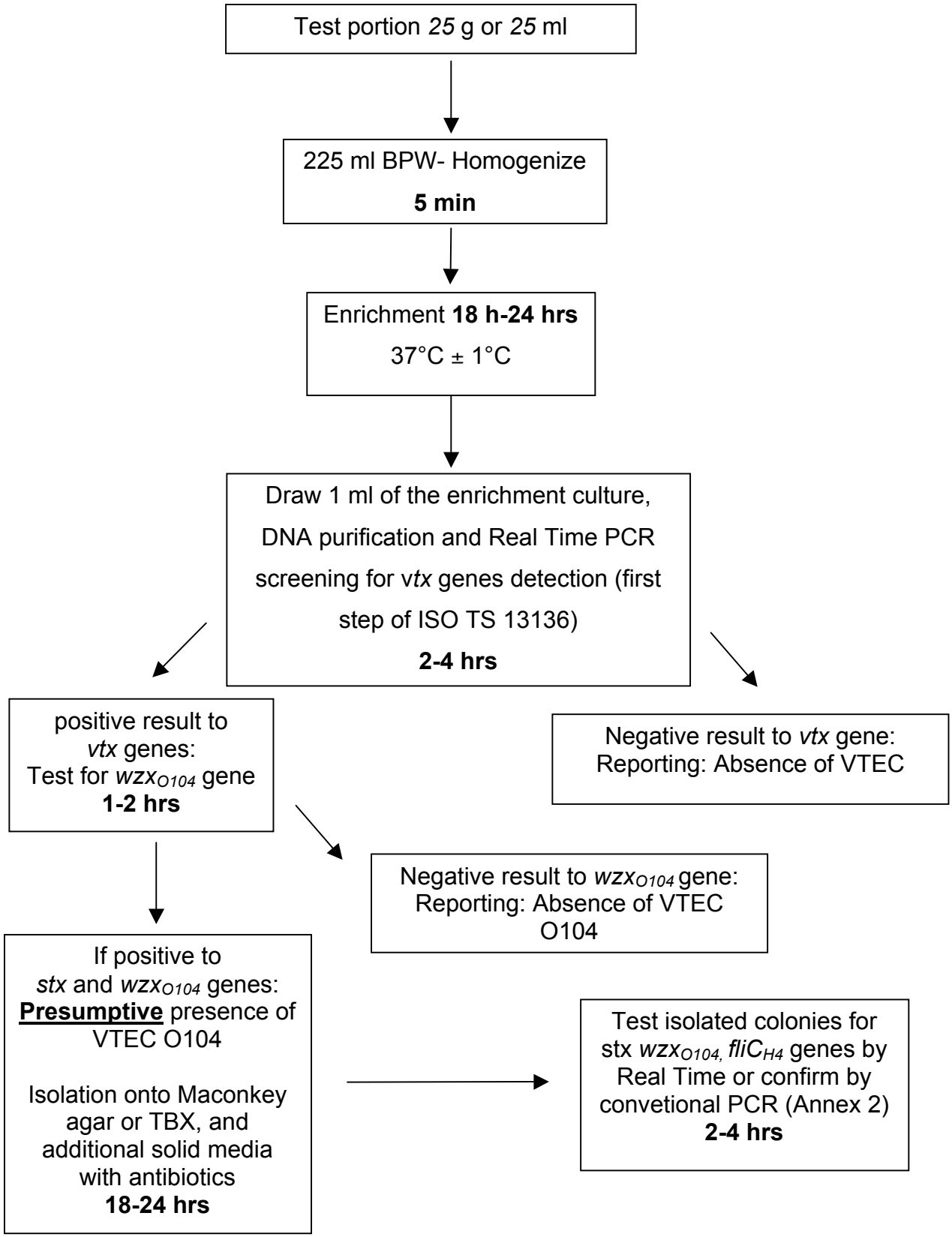
Enrichment cultures positive for presence of $wzx_{O104}$ gene are streaked onto MacConkey agar or TBX plates or any other media suitable for *E. coli* isolation. A second more selective plate can be chosen among the antibiotic-containing media described in the previous paragraph.

Up to 50 isolated colonies with typical *E. coli* morphology or growing on the media with antibiotics are collected and point-inoculated on Nutrient Agar (NA) (single colonies) and $H_2O$ (5 pools by 10 colonies each). *vtx* gene detection is performed on the isolated colonies or pools by Real Time or conventional PCR (reference methods can be found at http://www.iss.it/vtec). Colonies positive for *vtx* genes will be tested for the O104 antigen-associated gene $wzx_{O104}$ and the gene encoding the H4 flagellar antigen, $fliC_{H4}$.

For strain characterization, the Real-time PCRs for $wzx_{O104}$ and $fliC_{H4}$ can be run as duplex PCR, labeling the two probes with compatible phluorophors (e.g. FAM and HEX).

Alternatively, confirmation of isolated colonies as possessing the virulence profile of the German outbreak strain can be accomplished by using the conventional PCR protocol developed by the "Konsiliarlabor für Hämolytisch–Urämisches Syndrom (HUS)" and described in the "Laborinfo Stand 01.06.2011" downloadable from the website www.ehec.org.
The approach includes the detection of the *vtx, TerD,* $rfb_{O104}$, and $fliC_{H4}$ genes in a multiplex PCR reaction. A scheme of the procedure is in Annex 2.
As an optional step, the *vtx2* gene sub-typing may be carried out. The *vtx*-genotype of the German outbreak strain has been reported to be *vtx2a*. A conventional PCR for *vtx2*-genes subtyping has been distributed for the 6th proficiency test organised jointly by the EU RL VTEC in the framework of the 2010 work program and the WHO Collaborating Centre for Reference and Research on *Escherichia* and *Klebsiella*, at the Statens Serum Institute, Copenhagen, upon mandate of the ECDC.
An extract of the laboratory procedure sent for the 6th ring test describing the protocol for the detection of the *vtx2a* subtype is included in Annex 3.

**Flow-diagram of the screening procedure**



## Annex 1

## Real-time PCR for the detection of $wzx_{O104}$ and $fliC_{H4}$ genes

### 1. Principle of the method

This Real-time PCR protocol aims at the detection of the O104 serogroup-associated gene, $wzx_{O104}$, coding for the O-antigen flippase Wzx and the gene encoding the flagellar antigen H4, $fliC_{H4}$.

### 2. Operating procedure

The protocols are based on the 5' nuclease PCR assay. Considering that Real-time PCR may use different instruments and probes labeling chemistry, the amplification conditions to be applied may vary depending on the system used. Refer to the instructions supplied with the instrument and reagents' kit of choice.
A standard two-step thermal profile has been used at EU RL VTEC with satisfactory results for both the reactions. The details were as follows:

95°C X 10'

35 cycles of

95°C X 15''

60°C X 1'

The primers and probes to be used are listed in the table below. A typical Real Time PCR reaction is described below. The chemistry of the reporter and quencher phluorophores are not indicated being largely dependent on the Real-time PCR systems available in each laboratory.

For isolated strains characterization the Real-time PCR for $wzx_{O104}$ and $fliC_{H4}$ can be run as duplex PCR, labeling the two probes with compatible phluorophors (e.g. FAM and HEX).

RT PCR reaction assembly:

| | |
|---|---|
| Buffer 10X | to 1X ($MgCl_2$ 3mM) |
| Primer Fwd | 500 nM |
| Primer Rev | 500 nM |
| Probe | 200nM |
| DNA | 2 µl of DNA purified from 1 ml of culture and diluted 1:10 can be sufficient |
| Water | to final volume |

**Please note that due to the urgent necessity of the present protocol the procedure described above has been tested by the EU-RL VTEC on a limited set of isolated strains only and has not been validated on the enrichment cultures yet.**

### 3. Controls

A VTEC strain belonging to serotype O104:H4 should be used as positive control. DNA extracted by a VTEC strain belonging to serogroup O104 will be made available soon by the EU-RL VTEC to the NRLs which will require it.
The Real-time PCR procedure requires an **inhibition/extraction control**. Details on the possible systems to be used as inhibition/extraction control are given below. In particular, two different internal amplification controls (IACs) can alternatively be used:

- A commercially available TaqMan® Exogenous Internal Positive Control (Applied Biosystems, Foster City, CA, USA). The kit includes all reagents necessary (primers, a Vic™probe, IAC target DNA and blocking solution). The IAC target DNA must be diluted 10 times to achieve a copy number of approximately 100 per PCR reaction. The PCR product length is not declared to the customer.
- The open formula pUC 19 based internal amplification control IAC developed by Fricker et al. (2007. Appl Environ Microbiol 73, 1892-1898).

Approximately 100 copies of target DNA (pUC 19) should be used per PCR reaction. The size of the IAC is 119 bp.

The latter may be also used as an extraction control by adding 100 copies of the pUC 19 plasmid to the sample aliquot prior to the DNA purification step.

**Table 1. Primers and probes used in 5' nuclease PCR assays**

| Target gene (Ref.) | Forward primer, reverse primer and probe sequences (5'-3') [a] | Amplicon size (bp) | Location within sequence | GenBank accession number |
|---|---|---|---|---|
| *wzx*$_{O104}$(1) | TGTCGCGCAAAGAATTTCAAC<br>AAAATCCTTTAAACTATACGCCC<br>Probe-TTGGTTTTTTTGTATTAGCAATAAGTGGTGTC | 100 | 2,333,750–2,333,730<br>2,333,673–2,333,651<br>2,333,724–2,333,693 | CU928145 |
| *fliC*$_{H4}$(2) | GCTGGGGGTAAACAAGTCAA<br>CCAGTGCTTTTAACGGATCG<br>Probe- TCTTACACTGACACCGCGTC | 192 | 604-623<br>796-777<br>631-650 | AY249989 |

(1) Bugarel M. et al. Int J Food Microbiol 2010 142:318-329
(2) EU RL VTEC

## Annex 2

## Multiplex conventional PCR for the detection of *vtx, TerD,* $wzx_{O104}$ and $fliC_{H4}$ genes (From the "Laborinfo Stand 01.06.2011" developed at the "Konsiliarlabor für Hämolytisch– Urämisches Syndrom (HUS)" - www.ehec.org-).

**PCR zum spezifischen Nachweis von HUSEC041 (O104:H4)**

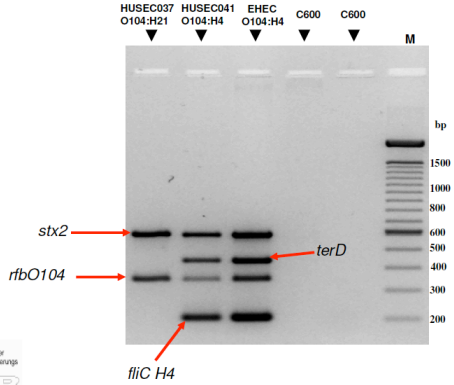
| Primerbezeichnung | Primersequenz (5'—3') | Target | Länge des PCR-Produkts (bp) | Referenz | Primerkonzentration |
|---|---|---|---|---|---|
| LP43 | ATC CTA TTC CCG GGA GTT TAC G | *stx2* | 584 | Cebula et al. 1995 | 150 nM |
| LP44 | GCG TCA TCG TAT ACA CAG GAG C | | | | |
| TerD1 | AGT AAA GCA GCT CCG TCA AT | *terD* | 434 | Taylor et al. 2002 | 240 nM |
| TerD2 | CCG AAC AGC ATG GCA GTC T | | | | |
| 104rfbO-f | TGA ACT GAT TTT TAG GAT GG | *rfbO104* | 351 | Hier | 360 nM |
| 104rfbO-r | AGA ACC TCA CTC AAA TTA TG | | | | |
| fliCH4-a | GGC GAA ACT GAC GGC TGC TG | *fliC H4* | 201 | Hier | 240 nM |
| fliCH4-b | GCA CCA ACA GTT ACC GCC GC | | | | |

**Reaktionsbedingungen**
25 µl PCR Ansatz mit REDTaq ReadyMix (Sigma-Aldrich)
mit 2 µl DNA Template

94 °C 5 min Denaturierung
30 Zyklen mit
1. Denaturierung 94 °C 30 s
2. Annealing 55 °C 60 s
3. Elongation 72 °C 60 s

72 °C 5 min. finale Elongation

**Nachweis im 2%igen Agarosegel (Exemplarisch in der Abbildung rechts)**




Alle Ergebnisse beziehen sich ausschließlich auf die untersuchten Proben. Ohne schriftliche Genehmigung des Prüflabors darf dieser Befundbericht nicht auszugsweise veröffentlicht werden.
***Parameter außerhalb des Akkreditierungsbereiches
Akkreditiert nach DIN EN ISO/EC 17025 2005 und erfüllt DIN EN ISO 15189 (2007)

DAC-ML-0174-02-10

## Annex 3

## Conventional PCR for the subtyping of the *vtx2* gene

### 1 Principle of the method

The method is based on PCR amplification of specific DNA regions from a template DNA, with oligonucleotides triggering the start of the PCR reaction. The procedure concerns the detection of the *vtx2a* gene subtype and is performed by specific PCR reactions, using primers designed on the basis of analyses of existing *vtx* sequences (reported in **Appendix 1**).

The method is composed of the following steps:

- Template preparation
- Setting up the PCR reaction
- Determination of the PCR results by agarose gel electrophoresis.

### 2 Template preparation

Isolated strains are streaked onto solid media (e.g. TSA) and incubated over night.

A single bacterial colony is inoculated in TSB and incubated over night.

25 µl of the overnight culture are added to 975 µl Milli Q water in Eppendorf tube and boiled for 15 minutes. Centrifuge at 18.000 g 5 minutes. Upon transfer to a clean tube, the supernatant is used directly for PCR and stored at -18ºC for further analyses.

### 3 Setting up the PCR reaction

PCR assays are set up in a total volume of 20 µl for standard PCR as follows:

1.25 µl $H_2O$
10 µl Mastermix (HotStart, Qiagen),
1.25 µl of each of three primers (STOCK solution of primers is 5 µM) §
5 µl supernatant of boiled lysate (STOCK)

**Primers' sequences:**

| | |
|---|---|
| vtx2a-F2 | GCGATACTG**R**G**B**ACTGTGGCC |
| vtx2a-R3 | CCG**K**CAACCTTCACTGTAAATGTG |
| vtx2a-R2 | GGCCACCTTCACTGTGAATGTG |

The thermal profile is:
95ºC for 15 min (HotStart Taq activation)
35 cycles of 94ºC for 50 sec, 64°C for 40 sec and 72ºC for 60 sec, ending with 72ºC for 3 min. PCR amplicons can be stored at 4°C until loading on agarose gel.

In each PCR assay, a positive and a negative control must be included. The DNA template to be used as positive control is available from the EU RL VTEC upon request. The negative control is constituted by a sample without template added.

**4 Agarose gel electrophoresis**

Prepare agarose gel in 1X Tris/Borate/EDTA (TBE) or Tris/Acetate/EDTA (TAE). Each well of the gel is loaded with 10 µl of each reaction added with loading dye at 1X final concentration. Run the samples in 1X running buffer (TBE or TAE) in constant voltage (100 V). Use a molecular weight marker suitable for assignment of the correct molecular weights to the amplicons produced (refer to Appendix 1). Consider that a correct band assignment is a crucial point in the assessment of the presence of the target genes. Make sure that the bands produced by the reference strains match exactly the expected molecular weight.

Agarose gels should be added of ethidium bromide to allow the visualisation of DNA. This reagent is a DNA intercalating agent commonly used as a nucleic acid stain in molecular biology laboratories. When exposed to ultraviolet light, it will fluoresce with a red-orange colour. Ethidium bromide should be added to a final concentration of 0.5 µg/ml before pouring the

agarose gel in the electrophoresis gel cast. Alternatively the agarose gel can be stained after electrophoresis in a 0.5 µg/ml ethidium bromide aqueous solution.

The expected amplicons size is about 350 bp.

## Annex 4

### Testing of seeds used for production of sprouts to be consumed raw

When the matrix to be analyzed is constituted by seeds the following considerations have to be made:

1. The seeds are generally contaminated at very low levels (0.1 to 1.8 cfu/gr as assessed for Salmonella. See reference below). Nonetheless, the sprouting process is characterized by conditions (humidity, heat) favoring the pathogen's enrichment.
2. The seeds are generally dried. Therefore, the contaminating pathogens are supposed to be stressed.
3. The contamination may occur on the surface of the seed as well as being internal. The latter case occurs when the seeds are contaminated during the growth of the plants used for their production (primary contamination of crops). Contamination of the seeds' surface can occur during all the phases of preparation, storage and general handling of the seeds (secondary contamination).
4. The enrichment cultures of seeds may contain inhibitors of the DNA polymerase used for the PCR screening of the samples

In order to increase the analytical power of the proposed method, the following steps are carried out when dealing with seeds:

1. 50 gr of seeds are analyzed instead of the usual 25 gr of food items in order to increase the sensitivity of the assay.
2. The seeds are smashed in a sterile container (e.g. a stomacher bag) using a mortar with pestel or other similar tools before adding the enrichment broth.
3. The smashed seeds are transferred to a sterile container (flask or a new stomacher bag) added with 450 ml BPW and incubated for 24 hrs at 37°C (either static or in agitation).
4. A 5 ml aliquot of the enrichment culture is taken, mixed by vortexing (in order to detach any possible Enteroaggregative bacteria adhering to seeds), centrifuged at 500 X g 1 min to sediment the seeds' debris.

One ml aliquot of the supernatant is taken at this stage and used for DNA preparation.

5. The DNA is diluted 1:10 before using it. In case of absence of amplification in the reactions containing the IAC, the DNA template is used at the dilution of 1:30.

Reference.

Ching-Hsing Liao, William F. Fett. 2003. Isolation of Salmonella from alfalfa seed and demonstration of impaired growth of heat-injured cells in seed homogenates. International Journal of Food Microbiology 82: 245– 253

# 6.Discussion

# 6. DISCUSSION

Verocytotoxin-producing *E. coli* are zoonotic pathogens able to cause severe disease in humans, such as haemorrhagic colitis and haemolytic uremic syndrome, and are a major issue in public health in industrialized countries.

The natural reservoir for these pathotype of *E. coli* is represented by ruminants, which are usually asymptomatic. Humans principally acquire the infection through the ingestion of contaminated meat, milk or other food contaminated by the ruminants' faecal matter, which do not undergo specific technological treatments able to reduce the bacterial load, such as pasteurization or sterilization. Infections can also be spread by environmentally related vehicles, such as drinkable or recreational water or vegetables contaminated after fertilization of fields with ruminants' manure and irrigative treatments with water from contaminated sources. Finally, the direct contact with infected animals during the practice of "petting zoo", when farms are visited for educational purposes is acquiring increasing importance (Heuvelink A.E. *et al.*, 2002).

Among the different food vehicles, VTEC infections related to the consumption of vegetables have recently acquired large mediatic attention, i.e. the occurrence of outbreaks linked to the ingestion of leaf vegetables in the US (Caprioli A. *et al.,* 2005; Wendel A.M. *et al.,* 2009) and the recent outbreak of infections occurred in Germany in 2011 associated to the consumption of sprouts.

VTEC are perhaps the most heterogeneous group of *E. coli.* Strains isolated so far from the animal reservoir belong indeed to more than 300 different serotypes. However, only a sub-population of these has been actually associated to human disease. O157 is the most common VTEC serogroup causing large outbreaks, characterized by a high proportion of severe cases compared to uncomplicated infections. Nevertheless, many sporadic cases and some of the largest outbreaks reported have been caused by non-O157 VTEC strains often including a high number of severe infections, such as the VTEC O145 that caused an outbreak in the US in 2010 with the romaine lettuce as the food vehicle, stressing the importance of defining the virulence potential of VTEC serogroups other than O157.

The pathogenicity of VTEC relies, beside the production of the verocytotoxins, on the presence of other virulence genes, all of them acquired by horizontal gene transfer. The most important accessory virulence determinant is represented by the LEE locus, which governs the ability of the strains to produce the *attaching-and-effacing* lesions on the host cell surface in the intestinal mucosa (see section 1.4.2), typically present in course of infection with the major VTEC serogroups.

Based on the phenotypic and molecular characteristics of the strains and on the clinical features of the associated diseases, VTEC have been categorized into groups termed seropathotypes (SPTs) named with letters from A to E (Table 1) according to a decreasing rank of pathogenicity. According to this classification, VTEC O157 composes the SPT A, while seropathotype B is made up of VTEC strains belonging to serogroups O26, O111, O103, O145 and O121, which are frequently isolated by severe disease in humans and, similarly to SPT A, are positive for the presence of the LEE locus. The SPTs C and D include VTEC strains, which are respectively rarely associated to HUS or causing uncomplicated

diarrhoea, while strains belonging to SPT E are supposed to circulate predominantly in the animal reservoir (Karmali M.A. *et al.*, 2003).

| Seropathotype | Relative incidence | Frequency of involvement in outbreaks | Association with severe disease[a] | Serotypes |
|---|---|---|---|---|
| A | High | Common | Yes | O157:H7, O157:NM |
| B | Moderate | Uncommon | Yes | O26:H11, O103:H2, O111:NM, O121:H19, O145:NM |
| C | Low | Rare | Yes | O91:H21, O104:H21, O113:H21; others |
| D | Low | Rare | No | Multiple |
| E | Nonhuman only | NA[b] | NA | Multiple |

[a] HUS or hemorrhagic colitis.
[b] NA, not applicable.

*Table 1: Classification of VTEC serotypes into seropathotypes (Karmali M.A. et al, 2003)*

The VTEC pathotype is thus divided into a number of sub-groups with differences in the genomic asset, the phenotypic traits as well as in the clinical outcome of the disease.

Moreover, large gaps still exist in the knowledge of the pathogenesis of the infections caused by VTEC, particularly by those serogroups negative for the presence of the LEE-locus, which are lacking of the main colonization mechanism and yet capable of causing severe disease.

These considerations set the need for conducting research studies for identifying the whole genomic asset needed for the full virulence of VTEC and for understanding the molecular bases of the VTEC-induced disease.

### 6.1 Identification and characterization of VTEC virulome: Characterization of a novel pathogenicity island encoding Subtilase cytotoxin

The production of the Subtilase cytotoxin (SubAB) has been described in LEE-negative VTEC strains. This toxin has been demonstrated to provoke damage in the unfolded protein response in the endoplasmic reticulum of the host cell, thus affecting the correct folding of proteins and finally inducing apoptosis. It has been proposed that SubAB could play a role in causing the disease through a synergistic activity with the VTs. SubAB has been first described in an O113 VTEC strain isolated during a small outbreak of HUS occurred in Australia in 2003 (Paton A.W. *et al.*, 2004). In the prototype strain the genes encoding the Subtilase (*subAB*) are vehiculated by a large virulence plasmid, named pO113, which also harbours the gene *saa*, encoding an autoagglutinating adhesin, which has been proposed to be involved in the colonization process (Paton A.W. *et al.*, 2001).

The production of SubAB has been also demonstrated in two *E. coli* strains, which did not belong to VTEC or AEEC, being at the same time negative for the presence of LEE locus and the Verocytotoxins-coding genes (strains ED 32 and ED 591). Such strains had been isolated in Italy from unrelated cases of non-complicated diarrhoea (Tozzoli R. *et al.*, 2010, Chapter 3). The *subAB* genes present in these strains were not identical to those described in the VTEC O113, but represented a different allelic variant sharing 90% of nucleotidic sequence similarity with the prototype operon and 93% of amino acidic identity (96% of positive substitutions) (Tozzoli R. *et al.*, 2010, Chapter 3). The isolates harbouring the *subAB* allelic

variant were also negative for the presence of *saa* and harboured the *subAB* operon on the chromosome, in a previously unidentified pathogenicity island that we have characterised in the genome of the *E. coli* strain ED 32 and termed Subtilase-Encoding PAI (SE-PAI) (Tozzoli R. *et al.*, 2010; Michelacci V. *et al.*, 2012; Chapter 3). This locus also contains other genes, which may have a role in the pathogenesis of the infections and are thus regarded as virulence genes. These include the *tia* gene, which encodes an invasion determinant described in ETEC strains associated to human disease (Fleckenstein J.M. *et al.*, 1996) and an homologous of the *shiA* gene (Michelacci V. *et al*., 2012, Chapter 3), which acts as an immunomodulator upon *Shigella flexneri* infections by attenuating the inflammatory response in the gut of the host (Ingersoll M.A. *et al.*, 2006). In the *E. coli* strain ED 32 the SE-PAI is located downstream the locus encoding *pheV*-tRNA (Michelacci V. *et al*., 2012, Chapter 3). This genomic area represents a hotspot for the integration of PAIs in different bacterial pathogens. As an example such a locus hosts the *she* PAI of *S. flexneri* 2a, (Al-Hasani K. *et al.*, 2001) and PAI $II_{AL862}$ in the extraintestinal pathogenic *E. coli* strain AL862 (Lalioui L. & Le Bouguenec C., 2001). Moreover, this locus represents the insertion site of the LEE PAI in many VTEC strains, including those belonging to serotype O103:H2, one of those most commonly associated with severe disease in humans (Jores J. *et al.*, 2001). It could be hypothesized that the LEE-negative, SE-PAI-positive VTEC may have arisen from a mutual exclusion event between this PAI and the LEE, which might have competed for the same integration site. This hypothesis could also explain in turn the apparent absence of the SE-PAI in LEE-positive strains.

The two *subAB* allelic variants show different behaviours in terms of geographic and epidemiologic distribution. Recent studies showed their differential distribution in VTEC strains circulating in different animal species. In particular, a predominance of the prototype subtilase-coding genes has been reported in LEE-negative VTEC strains isolated from cattle (named also $subAB_1$) (Orden J.A. *et al.*, 2011), although in presence of a relatively low prevalence. Conversely the allelic variant of *subAB*, also indicated as $subAB_2$, was prevalent in LEE-negative VTEC strains isolated from small ruminants, where the genes encoding this cytotoxin were generally highly represented (Orden J.A. *et al.*, 2011). These observations have been extended by investigating large panels of LEE-negative VTEC strains from healthy sheep and from human cases of diarrhoea isolated in Spain and Denmark respectively (Michelacci V. *et al*., 2012, Chapter 3). The study was conducted by using PCR tools specifically developed for the differential detection of the two *subAB* forms and led to the finding that *subAB* is also very well represented also in VTEC strains lacking the LEE locus and isolated from human cases of diarrhoea. In particular, the *subAB* genes were present in more than 70% of the isolates, thus supporting the hypothesis that subtilase cytotoxin may play a role in the pathogenesis of the disease caused by these strains (Paton A.W. *et al.*, 2004). To the best of our knowledge this represents the first direct evidence that *subAB* may be part of the LEE-negative VTEC virulome. Moreover, almost all the *subAB*-positive strains possessed the $subAB_2$ variant in combination with *tia* gene, suggesting that in these strains the Subtilase-coding genes were hosted into the SE-PAI.

Similar findings were observed in LEE-negative strains isolated from healthy sheep when assayed using the same strategy. The vast majority of the strains tested (86%) proved positive for *subAB* and all the strains possessed the $subAB_2$ allelic variant, mainly in association with the *tia* gene, suggesting the presence of SE-PAI.

Altogether, these findings seem to indicate that *subAB*$_2$ represents the prevalent allelic variant in the SubAB-producing, LEE-negative VTEC, at least in the strains circulating in Europe.
Moreover, sheep appear to be an important animal reservoir for the infections caused by *subAB*$_2$-positive and LEE-negative VTEC strain, including these strains among the *E. coli* strains with a zoonotic origin.

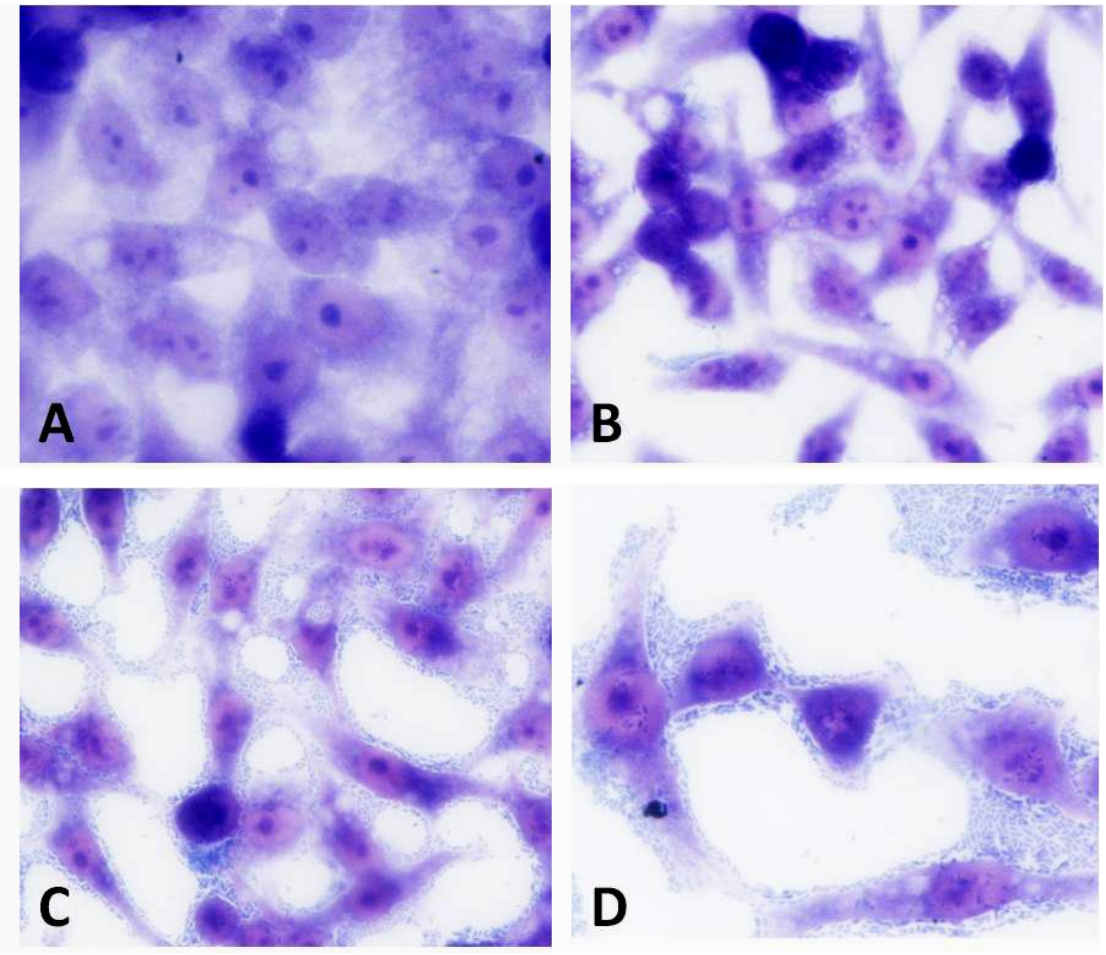



*Figure 5: Invasion assay on cultured Hep2 cells.* ***A)*** *Negative control: Hep2 cells not inoculated with E. coli strains;* ***B)*** *Hep2 cells upon inoculation with the E. coli K12 control strain MG1655;* ***C)*** *and* ***D)*** *Hep2 cells upon inoculation with the ED 32 strain.*

Despite the evidences in favour of a role of SE-PAI in causing the disease in humans, the function of the other virulence genes vehiculated by this PAI, such as the *tia* and *shiA* genes, still has to be determined. The product of *tia* gene has been demonstrated to be functional in the invasion of the host cells upon ETEC infections (Fleckenstein J.M. *et al*., 1996) and it could be hypothesized that it may have a similar role in the colonization process by the SE-PAI positive strains, with or without the concomitant action of the product of the co-localised *shiA* gene, acting as an alternative to the *attaching-and-effacing* machinery.
Preliminary experiments have been performed to assess the behaviour of SE-PAI-positive strains onto monolayers of cultured epithelial cells. The pattern of interaction of the ED 32 strain on cultured Hep-2 epithelial is shown in Figure 5. Interestingly, the ED 32 strain showed the ability to invade the host cells

(Figure 5, panels C and D): bacterial cells localize around the nucleus of the Hep-2 cells and not over them, distributing in the whole epithelial cells depicting their boundaries. Further studies are on-going to confirm the actual implication of the product of *tia* gene in the invasion of the host cells, by assaying the invasive ability of the ED 32 strain upon mutation of the *tia*.

**6.2 Identification and characterization of VTEC virulome: the Genomic Island (GI) OI-57.**

The advent of whole-genome probing and sequencing technologies substantially enhanced the identification of bacterial genomic regions playing a role in the virulence. Such high-throughput approaches have been effective in studying the VTEC genomic structure allowing the identification of new virulence-associated genetic elements.

VTEC strains belonging to serogroups associated with SPTs A and B (Table 1) are considered to be the most pathogenic VTEC, being associated with the induction of the most severe form of the infection, HUS. These isolates possess the LEE-locus and the genomic island named OI-122 in the genome of EDL933 VTEC O157:H7 reference strain (Morabito S. *et al.*, 2003). However, their complete asset of virulence genes still needs to be completely defined. This was the aim of a study, object of the paper described in chapter four, based on the whole genome comparison of strains belonging to the different SPTs by microarray hybridization. This approach allowed the identification of a GI, termed OI-57 in the chromosome of the VTEC O157 EDL933 strain, consistently present in VTEC strains belonging to SPTs A and B and not in the less pathogenic strains part of SPT C, SPTs D and E (Imamovic L. *et al.*, 2010, Chapter 4).

The OI-57 harbours two adjacent genes whose products are annotated with a function: *adfO*, encoding a putative virulence factor promoting the adherence of VTEC O157 to HeLa cells (Ho T.D. *et al.*, 2008), and *ckf*, encoding a phage-associated putative bacterial cell-killing factor (Perna N.T. *et al.*, 2001; Imamovic L. *et al.*, 2010; Chapter 4).

The presence of OI-57 has been observed also in EPEC strains, either typical or atypical. This finding is not surprising: tEPEC share with VTEC belonging to SPTs A and B the ability to cause the *attaching-and-effacing* lesion to the host intestinal mucosa, while aEPEC, in addition to the LEE locus, have in common with these VTEC also other parts of their genomic background such as the large virulence plasmid harbouring the enterohaemolysin-coding operon. Altogether, these observations extend to the OI-57 the spectrum of mobile genetic elements (MGEs) shared by VTEC and EPEC (sometimes indicated as the unique group of Attaching-and-Effacing *Escherichia Coli*, AEEC) and suggest a possible involvement of the genes vehiculated by OI-57 in the enterocyte colonization processes *via* the A/E mechanism. The hypothesis of an involvement of OI-57 in the colonization of the host gut is strengthened by the finding that the island is entire in all the VTEC strains analysed and that the two genes putatively involved in the attachment are also present in the most the other AEEC strains analysed although with a certain degree of variability in the other regions of the genomic island (Imamovic L. *et al.*, 2010, Chapter 4).

OI-57 appears thus to be part of the virulome of VTEC strains causing severe disease in humans and seems to be crucial in the induction of the *attaching-and-effacing* colonization of the intestinal mucosa.

Further studies are needed to assess the effective role of *adfO* and *ckf* as virulence factors during the pathogenesis of the disease.

Furthermore, due to the strong association with VTEC belonging to SPT A and B, OI-57 genes can be considered as good candidates to be virulence predictors and potentially useful as targets for the development of diagnostic tools aimed at the identification of these VTEC SPTs.

The currently used method for the detection of pathogenic VTEC in food matrices (ISO/TS 13136:2012) includes the Real-Time PCR screening of enrichment cultures for the presence of genes encoding VT production (*vtx*). The samples positive at the screening for *vtx* are subjected to isolation. Enrichment cultures are also screened for the presence of the *eae* gene, encoding the intimin adhesin, responsible for the *attaching-and-effacing* lesion, as well as of genes associated to the main VTEC serogroups causing severe human disease (O157, O26, O103, O145, O111 and O104). This method has been recently approved and has the advantage of returning negative results in just one day, compared to the three days required for classical cultural methods.

The development of even faster methods for the detection of the most pathogenic VTEC seropathotypes in food matrices and biological samples is the object of an existing collaboration between the Istituto Superiore di Sanità (ISS, Rome) and BioRad-Food Science Division (Marne La Coquette, Paris). In the framework of this project, three open reading frames carried by the OI-57 (named Z2096, Z2097 and Z2121 in EDL933 reference genome, Acc. No. NC_002655) have been considered as good candidates to be used as targets in molecular assays specific for the detection of the most pathogenic VTEC. In fact, in preliminary standard PCR screenings the three of them resulted present in all the assayed strains belonging to SPT B, in most of the SPT A and in a few VTEC belonging to SPT C, and completely absent in all the strains belonging to SPTs D and E and the *E. coli* from the reference ECOR collection (Imamovic L. *et al.*, 2010, Chapter 4). These genes have been evaluated as single targets for the detection of the most pathogenic VTEC in food matrices using nucleic acids-based methods. We have developed a Real-Time PCR assay specific for each of the three OI-57 targets (Z2096, Z2097 and Z2121) and tested their use on a large panel of *E. coli* strains belonging to different pathotypes in order to assess their exclusivity and inclusivity. In particular, the three ORFs were present in all the VTEC strains belonging to SPTs A and B (41/41), a few SPT C (Z2096 in 5/12 and Z2097 in 2/12 strains) and in some EPEC strains (Z2096 and Z2097: 14/42; Z2121: 24/42).

| | | **Tot. No. of strains** | **No. of strains positive for** | |
|---|---|---|---|---|
| | | | **Z2096** | **Z2096$_{VTEC}$** |
| **EPEC** | | 42 | 14 | 4 |
| **VTEC** | **SPT A** | 9 | 7 | 7 |
| | **SPT B** | 40 | 40 | 27 |
| | **SPT C** | 12 | 5 | 0 |
| | **SPT D/E** | 6 | 0 | 0 |

*Table 2: Results from the screening of the EPEC and VTEC panel for the presence of ORF Z2096.*

The sequences of the three genes, amplified from the positive EPEC strains, were compared to the corresponding sequences of VTEC strains available in GenBank, in order to identify polymorphisms enabling the development of primers and probes to be used in Real-Time PCR assays specifically targeting the VTEC-associated forms of the genes. A single nucleotide polymorphism has been detected in the VTEC-associated form of the ORF Z2096 and a primer pair specific for the VTEC variant of Z2096 has been developed and tested on a large panel of VTEC, EPEC and ECOR control strains (Table 2).

This approach resulted in the elimination of the interference from the strains belonging to SPT C and in the reduction of positivity by EPEC (Table 2). However, 13 out of the 40 strains belonging to SPT B failed to be amplified with the specific primers for the VTEC variant (Table 2). This finding could be explained with the existence of another allele present in the negative SPT B VTEC or the presence, in these strains, of the EPEC-associated variant. The latter scenario could be possible, given the wide genomic overlap between VTEC and aEPEC. Indeed, whether aEPEC and VTEC can be considered as two distinct classes of pathogenic *E. coli* or the latter being VTEC that lost the *vtx*-phage is still under debate.

In conclusion, the Real-Time PCR diagnostic assay developed is able to correctly detect VTEC strains belonging to the seropathotypes A and B, reported to be more frequently associated with cases of severe disease in humans. Nevertheless, while giving excellent results concerning the exclusivity, the parameter of inclusivity of such a strategy still needs to be implemented. An option for the optimization of this technique, based on the inclusion of an additional target, has been considered and is now under evaluation. Preliminary results showed that such a target could be represented by *toxB* gene, present on the large virulence plasmid characterising the SPT A and B VTEC, which appears to be present in all the strains belonging to SPTs A and in many of the SPT B VTEC. Such an additional target could efficiently complement the performance of the $Z2096_{VTEC}$ assay by increasing the inclusivity to all the SPT B VTEC isolates (Manuscript in preparation).

If proved successful, such a strategy could allow the development of a diagnostic tool for the detection of the most pathogenic VTEC in any matrix regardless the serogroup, representing a substantial improvement of an approach that has demonstrated its weakness during the last VTEC O104:H4 large outbreak occurred in Germany and France last may and June 2011. Moreover, despite the inclusion of an additional gene in the Real-Time PCR screening, *toxB,* the strategy would remain "single target", since the necessary condition for a sample to be considered as positive would be the presence of one among the targets tested. A further improvement of the strategy could be the use of cDNA preparations as templates or intercalating agents such as ethidium or propidium mononazide during the DNA extraction phase. The latter compounds would prevent the free DNA and nucleic acid deriving from damaged and dead cells to be amplified allowing the direct assessment of the viability of the detected VTEC strain without proceeding to the isolation using microbiological methods.

### 6.3 VTEC strains possessing rare combinations of MGEs vehiculating virulence factors: the O104 example

The extraordinary genomic plasticity of *E. coli* favours the continuous remodelling of the genome through the horizontal gene transfer. Such genomic dynamism accounts for the selection of chimeric pathotypes. This is what it has been proposed for the emergence of the VTEC O104 that caused the large outbreak of infections in Germany and France in May and June 2011. The causative agent was quite unusual as a VTEC. It belonged to a rare serotype, O104:H4, and lacked the main accessory virulence-associated gene, *eae*, considered the hallmark of highly pathogenic VTEC. Genomic analysis have shown that this strain shared the genomic backbone of a typical Enteroaggregative *E. coli* strain (EAEC, see section 1.4.2), that acquired the ability to produce verocytotoxins following the acquisition of a VT-coding bacteriophage (Brzuszkiewicz E. *et al.*, 2011; Holger Rohde M.D. *et al.*, 2011).

The episode was amongst the largest VTEC outbreaks ever recorded, involving some 4,000 cases of infection, and undoubtedly the most severe, counting an unusually high proportion of patients affected by haemolytic uremic syndrome.

Before the German outbreak, such an atypical VTEC had caused a few sporadic cases of infection in the EU in the past ten years. In particular one of these strains was retrospectively identified as the cause of an HUS case occurred in Italy in 2009. The patient was a child, which was affected by bloody diarrhoea, followed by HUS and neurological symptoms, showing long-term sequelae such as stage IV chronic kidney disease, hypertension and severe visual impairment (Scavia G. *et al.*, 2011, Chapter 5). The strain showed a PFGE-profile with a high degree of similarity (94,7%) with the strain responsible for the German outbreak. An epidemiological interview helped to trace back the acquisition of the infection to a recent travel in Tunisia. On the other hand the VTEC O104 that caused the German outbreak has been linked to eating sprouts produced by fenugreek seeds grown in Egypt and distributed in Germany and other European countries (Technical report of EFSA, July 5$^{th}$ 2011). Moreover, a connection with travel in North African countries was also assessed for at least one of the other sporadic cases (Finland 2010: Scientific report of EFSA, October 3$^{rd}$ 2011).

The provenance of these atypical VTEC from developing countries may explain their emergence. In these parts of the world, where the burden of disease caused by diarrheagenic *E. coli*, including EAEC, is very high, it is conceivable that an high load of such pathogens is released in the ecosystem and particularly in the aquatic environment. Indeed wastewater treatment is rarely implemented in these countries and water represents the main source of enteric infection for humans and it is not infrequent that humans and animals, particularly ruminants, share the environment in a promiscuous life style (Figure 4). Such a scenario describes the best setting for the two natural reservoirs for both EAEC and VTEC to be kept together, providing an explanation for the occurrence of a contact between the two *E. coli* pathotypes and the acquisition, by an EAEC, of the VT-coding phages by horizontal gene transfer.

The emergence of new chimeric pathotypes of *E. coli*, together with the possibility that human-borne pathogens from developing countries may jump into the food chain in the industrialized areas of the globe, raises questions concerning the perspective of controlling the food in order to prevent humans from acquiring the infection. As an example, when the German outbreak occurred in 2011 no methods were available for the specific detection of the outbreak strain.

With the aim of providing the laboratories involved in diagnostic analysis and food testing during the epidemiological surveys with the necessary tools, we have developed Real-Time PCR reagents specific for the detection of the O104 and H4 antigen-associated genes (*wzxO104* and *fliCH4*). Such reagents have been deployed using either published (Bugarel M. *et al.*, 2010) or directly determined sequences (*fliCH4*, Acc. No. AY249989) and had the goal of being compatible with those already used for the detection of the *vtx* genes. A dedicated laboratory procedure was produced and published on the EU RL VTEC website (EU Reference Laboratory for. *E. coli*. 2011, Chapter 5; http://www.iss.it/binary/vtec/cont/Lab_proc_O104_rev2.pdf).
The proposed method represents an important tool for the fast and specific detection of this rare and unusual VTEC serotype in foodstuff and, during the German outbreak, it largely contributed to the laboratory analyses performed on food samples aimed at identifying the vehicle of the infection.

**6.4 Concluding remarks**

VTEC infections represent a significant public health issue worldwide and, at the same time, their management is complicated by the large heterogeneity of the pathogroup either at genomic and phenotypic level. The complete genomic asset conferring virulence to VTEC has not been completely determined and neither are the complete pathogenetic mechanism or the epidemiology of the infections. This has reflections on both the prevention and clinical management of infections.
The results obtained in this piece of research addressed some of the open questions in the field of VTEC infections. In particular, two new pathogenicity Islands have been identified as part of VTEC virulome and characterised:

- The Subtilase-Encoding PAI (SE-PAI) vehiculates a newly described allelic variant of the Subtilase cytotoxin in association with two other virulence genes probably involved in the colonization mechanism: *tia* and *shiA*.

- The OI-57, present in VTEC serogroups belonging to the most pathogenic seropathotypes, harbours two putative virulence genes: *adfO*, encoding an adhesion factor, and *ckf*, annotated as a phage-associated bacterial cell-killing factor.

The analysis of the distribution and the function of the genes carried by these PAIs led us to make hypotheses on the pathogenetic mechanism of VTEC infections:

- The SE-PAI appears to contribute to induction of the disease caused by some VTEC clones. In fact it was detected in the vast majority of LEE-negative VTEC strains isolated from human cases of infections. This PAI may act by either producing the Subtilase cytotoxin, which exerts a cytopathic effect by affecting the correct folding of proteins in the host cells, or encoding colonization factors. In particular, the product of *tia* gene is almost identical to a factor involved in the invasion of the host cells and produced by ETEC, while the protein encoded by the *shiA*, resembles the product of the homologous gene of *Shigella flexeneri*, which reduces the inflammatory response upon infection. These additional virulence factors could be involved in the mechanism used

by LEE-negative VTEC to efficiently colonize the host in alternative to the *attaching-and-effacing* lesion.

- The products of the virulence genes carried by the OI-57 seem to take part to the complex *attaching-and-effacing* hystopathological lesion. In particular the gene *adfO*, whose product has been described as a factor that enhances the capability of the VTEC O157 strains to adhere to monolayers of cultured HeLa cells (Ho T.D. *et al.*, 2008). The finding, obtained in the framework of the work presented here, that the OI-57 is present and entire in all the LEE-positive VTEC and that *adfO* gene is carried by the most of the EPEC strains strengthens this hypothesis.

The study of the genomic asset and the distribution of the virulence-associated mobile genetic elements allowed us to make some inferences on the provenance of particular VTEC clones providing the basis for further epidemiological investigations on the human infections.

- The massive presence of the SE-PAI in LEE-negative VTEC either isolated from human infection or circulating in sheep indicates this species as the probable animal reservoir for these strains, suggesting that food of ovine origin might represent the main vehicle for the spreading of human infections and indicating a zoonotic origin for these pathogens.

- The results from the studies on the VT-producing EAEC (VT-EAEC) strains provided insights on the provenance of this aggressive chimeric *E. coli* pathotype. The connection with low-income countries of the VT-EAEC O104 that caused the outbreak in Germany in 2011 and of the other VT-EAEC O104 strains sporadically isolated in the EU in the last 11 years sets these geographic areas as the most probable reservoir for these peculiar VTEC strains. This hypothesis is supported by the genomic stability of the clone, as demonstrated by the analysis conducted by comparing the VTEC O104 that caused the German outbreak with those isolated from the HUS cases occurred in Germany in 2001 and in Italy in 2009. The results of this study suggest that the clone emerged and established in developing countries, probably in the North African area, and has been introduced into the EU several times from its first appearance in 2001 (Mellmann A. *et al.*, 2008).

The natural abundance of mobile genetic elements carrying virulence-associated genes, in combination with the extraordinary plasticity of *E. coli* genome, opens the way for the continuous emergence of new pathotypes. When such phenomena occur, no canonical categorization is sufficient to classify the emerging strains, without taking into account the overall characteristics of the strains. This was the case of the O104:H4 VT-producing EAEC strain responsible for the German outbreak of infections in 2011, which was at first considered as a typical VTEC strain belonging to an unusual serotype. After virulotyping and whole-genome sequence analysis the strain was then correctly classified as an EAEC

strain, which had acquired the bacteriophage vehiculating the verocytotoxin-coding genes by horizontal gene transfer.

Food and clinical microbiologists should therefore become more familiar with the virulotyping concept to classify *E. coli* according to their asset of virulence determinants, while keeping additional information, such as the serotype they belong to, as markers for the epidemiological investigation during outbreaks.

**6.5 References**


- Al-Hasani K, Rajakumar K, Bulach D *et al*. 2001. **Genetic organization of the she pathogenicity island in *Shigella flexneri* 2a.** Microb. Pathog. 30:1-8.
- Brzuszkiewicz E, Thürmer A, Schuldes J *et al.* 2011. **Genome sequence analyses of two isolates from the recent *Escherichia coli* outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-Haemorrhagic *Escherichia coli* (EAHEC).** Arch Microbiol. 193(12):883-91.
- Bugarel M, Beutin L, Martin A. *et al*. 2010. **Micro-array for the identification of Shiga toxin-producing *Escherichia coli* (STEC) seropathotypes associated with Hemorrhagic Colitis and Hemolytic Uremic Syndrome in humans.** Int J Food Microbiol. 142(3):318-29.
- Caprioli, A., Morabito, S., Brugère, H., Oswald, E. 2005. **Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission.** Vet. Res. 36: 289–311.
- EU Reference Laboratory for. *E. coli*. 2011. **Detection and identifi cation of Verocytotoxin-producing *Escherichia coli* (VTEC) O104:H4 in food by real time PCR—laboratory procedure**. http://www.iss.it/binary/vtec/cont/Lab_proc_O104_rev2.pdf
- Fleckenstein JM, Kopecko DJ, Warren RL, Elsinghorst EA. 1996. **Molecular characterization of the *tia* invasion locus from enterotoxigenic *Escherichia coli*.** Infect Immun. 64(6):2256-65.
- Heuvelink AE, van Heerwaarden C, Zwartkruis-Nahuis JT *et al*. 2002 ***Escherichia coli* O157 infection associated with a petting zoo.** Epidemiol. Infect. 129: 295-302.
- Ho TD, Davis BM, Ritchie JM, Waldor MK. 2008. **Type 2 secretion promotes enterohemorrhagic *Escherichia coli* adherence and intestinal colonization.** Infect. Immun. 76:1858–1865.
- Holger R, Qin J, Cui Y *et al.* and the *E. coli* O104:H4 Genome Analysis Crowd-Sourcing Consortium. 2011. **Open-Source Genomic Analysis of Shiga-Toxin–Producing *E. coli* O104:H4**. N Engl J Med. 365:718-724.
- Imamovic L, Tozzoli R, Michelacci V *et al*. 2010. **OI-57, a genomic island of *Escherichia coli* O157, is present in other seropathotypes of Shiga toxin-producing *E. coli* associated with severe human disease.** Infect Immun. 78(11):4697-704.
- Ingersoll MA, Zychlinsky A. 2006. **ShiA abrogates the innate T-cell response to *Shigella flexneri* infection.** Infect. Immun. 74:2317-2327.
- Jores J, Rumer L, Kiessling S, Kaper JB, Wieler LH. 2001. **A novel locus of enterocyte effacement (LEE) pathogenicity island inserted at *pheV* in bovine Shiga toxin-producing *Escherichia coli* strain O103:H2.** FEMS Microbiol Lett. 204:75-79.
- Karmali MA, Mascarenhas M, Shen S *et al*.. 2003. **Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease**. J Clin Microbiol. 41(11):4930-40.
- Lalioui, L., Le Bouguenec, C. 2001. **afa-8 Gene cluster is carried by a pathogenicity island inserted into the tRNA(Phe) of human and bovine pathogenic *Escherichia coli* isolates.** Infect. Immun. 69, 937–948.

- Mellmann A, Bielaszewska M, Köck R *et al*. 2008. **Analysis of collection of hemolytic uremic syndrome-associated enterohemorrhagic *Escherichia coli*.** Emerg Infect Dis. 14(8):1287-90.
- Michelacci V, Tozzoli R, Caprioli A *et al.* 2012. **A new pathogenicity island carrying an allelic variant of the Subtilase cytotoxin is common among Shiga toxin producing *Escherichia coli* of human and ovine origin.** Clin Microbiol Infect. Accepted for publication.
- Morabito S, Tozzoli R, Oswald E, Caprioli A. 2003. **A mosaic pathogenicity island made up of the locus of enterocyte effacement and a pathogenicity island of *Escherichia coli* O157:H7 is frequently present in attaching and effacing *E. coli*.** Infect Immun. 71(6):3343-8.
- Orden JA, Horcajo P, de la Fuente R *et al*. 2011. **Subtilase cytotoxin-coding genes in verotoxin-producing *Escherichia coli* strains from sheep and goats differ from those from cattle.** Appl Environ Microbiol. 77(23):8259-64.
- Paton AW, Srimanote P, Talbot UM *et al*. 2004. **A new family of potent AB(5) cytotoxins produced by Shiga toxigenic *Escherichia coli*.** J Exp Med. 200:35–46.
- Paton AW, Srimanote P, Woodrow MC, Paton JC. 2001. **Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans.** Infect Immun. 69(11):6999-7009.
- Perna NT, Plunkett G 3rd, Burland V *et al*. 2001. **Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7**. Nature. 409: 529-533.
- Scavia G, Morabito S, Tozzoli R *et al*. 2011. **Similarity of Shiga toxin-producing *Escherichia coli* O104:H4 strains from Italy and Germany**. Emerg Infect Dis. 17(10):1957-8.
- SCIENTIFIC REPORT OF EFSA. **Shiga toxin-producing *E. coli* (STEC) O104:H4 2011 outbreaks in Europe: Taking Stock**, European Food Safety Authority, October 3rd 2011; http://www.efsa.europa.eu/en/efsajournal/doc/2390.pdf
- TECHNICAL REPORT OF EFSA. **Tracing seeds, in particular fenugreek (*Trigonella foenum-graecum*) seeds, in relation to the Shiga toxin-producing *E. coli* (STEC) O104:H4 2011 Outbreaks in Germany and France**, European Food Safety Authority, July 5th 2011; http://www.efsa.europa.eu/en/supporting/doc/176e.pdf
- Tozzoli R, Caprioli A, Cappannella S *et al*. 2010. **Production of the subtilase $AB_5$ cytotoxin by Shiga toxin-negative *Escherichia coli*.** J Clin Microbiol. 48(1):178-83.
- Wendel AM, Johnson DH, Sharapov U *et al*. 2009. **Multistate outbreak of *Escherichia coli* O157:H7 infection associated with consumption of packaged spinach, August-September 2006: the Wisconsin investigation.** Clin. Infect. Dis. 48(8):1079-86.

# 7.Appendices

# Appendix I: List of abbreviations

**A/E:** *attaching-and-effacing*
**DAEC:** diffusely adherent *E. coli*
**DEC:** diarrhoeagenic *E. coli*
**bp:** base pair
**Da:** Dalton
**DNA:** deoxyribonucleic acid
**ds:** duble-stranded
**EAEC:** enteroaggregative *E. coli*
**EIEC:** enteroinvasive *E. coli*
**EHEC:** enterohaemorrhagic *E. coli*
**EPEC:** enteropathogenic *E. coli*
**ETEC:** enterotoxigenic *E. coli*
**ExPEC:** extraintestinal pathogenic *E. coli*
**GI:** genomic island
**HC:** haemorrhagic colitis
**HUS:** haemolytic uraemic syndrome
**kb:** kilobase
**LEE:** locus of enterocyte effacement
**LT:** heat-labile toxin
**MGE:** mobile genetic element
**NMEC:** neonatal meningitis *E. coli*
**ORF:** Open reading frame
**PAI:** pathogenicity island
**ST:** heat-stable toxin
**VT1:** Verocytotoxin 1
**VT2:** Verocytotoxin 2
**STEC:** Shiga toxin-producing *E. coli*
**Tir:** translocated intimin receptor
**T3SS:** type III secretion system
**UPEC:** uropathogenic *E. coli*
**VTEC:** Verocytotoxin-producing *E. coli*

# Appendix II: List of figures and tables

# Appendix III: Acceptance letter for Publication n.2 presented in Chapter 3

**From:** <q001fa@i-med.ac.at>
**Date:** 02 dicembre 2012 11:09:59 CET
**To:** <stefano.morabito@iss.it>
**Cc:** <q001fa@i-med.ac.at>, <franz.allerberger@ages.at>
**Subject: Acceptance for publication, Clinical Microbiology and Infection**

Clinical Microbiology and Infection
Official Publication of the European Society of Clinical Microbiology and Infectious Diseases

02-Dec-2012

CLM-12-5122.R1
A new pathogenicity island carrying an allelic variant of the Subtilase cytotoxin is common among Shiga toxin producing Escherichia coli of human and ovine origin
Michelacci, Valeria; Tozzoli, Rosangela; Caprioli, Alfredo; Martinez, Remigio; Scheutz, Flemming; Grande, Laura; Sanchez, Sergio; Morabito, Stefano

Dear Dr. Stefano Morabito

Your submission has now been provisionally accepted for publication in Clinical Microbiology and Infection (Original Article ).

The manuscript will be edited prior to creation of proof pages and you will be contacted if there are author queries that require your attention. If not, you will have an opportunity to review the edited manuscript at the stage of proofs.

Please see important details concerning the publication process below.

Your efforts during the process of revision are acknowledged and I hope you also are pleased with the final result.

We appreciate being able to publish your work and look forward to seeing your paper online as soon as possible.

Best regards
Prof. Franz Allerberger
Editor
Clinical Microbiology and Infection