ROMA TRE
UNIVERSITÀ DEGLI STUDI

DOCTORAL SCHOOL IN BIOLOGY

Section: Biology Applied to Human Health

CYCLE XXVII

Evolution of virulence in pathogenic Escherichia coli strains: Impact on public health

Laura Grande,
MSc. (Hons), Biology

Tutor
Prof. Paolo Visca

Co-Tutor
Dr. Stefano Morabito

PhD Coordinator
Prof. Paolo Visca

Abstract

Diarrheagenic *Escherichia coli* (DEC), including Verocytotoxin-producing *E. coli* (VTEC), are a significant public health issue worldwide. The management of the infections caused by these bacterial pathogens is complicated by their extreme heterogeneity, including strains causing a plethora of symptoms spanning from uncomplicated diarrhoea to life-threatening systemic sequelae such as the haemolytic-uremic syndrome (HUS).

Considerable efforts have been devoted by the scientific community to understand the evolutionary forces leading to the emergence of new *E. coli* pathogenic clones. However, *E. coli* genome is very dynamic and evolves continuously through horizontal gene transfer. The main objective of the present piece of research was to investigate the molecular bases of the evolution of pathogenic *E. coli,* mainly DEC.

We focused on VTEC as a model of dynamic pathogenic group (pathotype) encompassing many established *Escherichia coli* pathotypes. VTEC pathogenicity mainly relies on two aspects, the ability to efficiently colonize the host gut and the production of the Verocytotoxins (VTs). Genes encoding the VTs (*vtx*) are carried by lambdoid bacteriophages whose genome is normally integrated into the *E. coli* chromosome.

We have investigated the distribution of several Mobile Genetic Elements (MGEs) encoding virulence features in different VTEC subpopulations and studied their evolution through the differentiation of the virulence genes into different allelic variants. Additionally, we studied the different bacteriophages transporting the *vtx* genes into different VTEC types and their role in manipulating the host biology. Finally, we assessed the possibility that the spreading of VT-phages may or may not be subjected to host-related barriers by probing the possible acquisition of such phages by *E. coli* strains with a genetic background different from VTEC.

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. This approach led to the identification of a pathogenicity island encoding an allelic variant of the Subtilase cytotoxin (*subAB2*, Chapter 3) and to the description of the allelic variants of the colonization-associated virulence factor *toxB* (*toxB1* and *toxB2*, Chapter 4). Moreover, the distribution analysis of the two virulence factors on a large panel of *E. coli* strains allowed making inference on their role in the VTEC pathogenetic process.

The genomic approach has also been used to identify and characterize different VT-phages present in *E. coli* strains associated to the most severe form of infection, the HUS. In detail, a VT-phage present in VTEC O157 strains and able to influence the regulation of genes involved in the colonization mechanism has been described (Φ-8, Chapter 5). Additionally, the complete sequence of the VT-phage isolated from a recently described *E. coli* pathotype, the Enteroaggregative Haemorrhagic *E. coli* (EAHEC) has been obtained and compared to other VT-phages sequences with the aim to help unravelling their complicate biology (Phi-191, Chapter 5). With the same aim, the evaluation of the stable acquisition of VT-phages by non and pathogenic *E. coli* strains belonging to all the known pathotypes led to the conclusion that such phages show a host range broader than expected (Chapter 6). This observation, together with other evidences from the literature raises the hypothesis that probably any Enterobacteria equipped with an efficient colonization machinery could be potentially acquire a VT-phage generating clones with augmented virulence potential for humans, as it happened with the EAHEC O104:H4 that caused a large outbreak of HUS in Germany in 2011.

Acknowledgements

Looking back, I can confirm this is the most difficult part of my PhD thesis to be written.

This piece of Research is the result of a three-years period spent at the Unit of Foodborne Zoonoses at the Istituto Superiore di Sanità (ISS), where periods of pressure were often alternated by pleasant moments. I'm really grateful and at the same time conscious I could never hope for myself a better place to spend this period, from both a human and a scientific side.

Many thanks to prof. Paolo Visca, who acted as a supervisor before and during my PhD program.

My warm thanks goes to dr. Alfredo Caprioli, first of all for guesting me at the Unit he heads at ISS, for always being encouraging during this period and always listening to me whenever I needed, no matter how busy he could be.

I owe much gratitude to my supervisor dr. Stefano Morabito, who patiently tried to teach me how to move in the fascinating world of Research and for always encouraging me to give the best. I also thank him for his passionate attitude in work and life, for always being a enjoyable and optimistic person, and for often representing a friend, other than a supervisor. I just hope to make him proud of his investment on me...

I would really want to warmly thank the "droplets", Valeria, Rosy, and Antonella. Vale, for taking care of me in my really first steps in a lab, for being a strong and precious support even in my worst moments, for our gossip time, both hilarious and serious. Rosy, thanks for being first of all a guide and a great example to follow, an advisor, a precious friend, depending on my needs. Anto, for her freshness and spontaneity, for often being a "life coach" for all of us, for her special way to help me in my toughest moments, and for the tears she will shed while reading these words.

I'm also very grateful to all the people I met during this period, students, trainees and lab guests, they unconsciously and always spur on me to be better. Among them, a special thanks goes to Paola, for her extraordinary sensitivity, to Fernanda, for always making me laugh, to Federica, for her hilarious tales.

A really important step for my education was the possibility I had to go abroad and spend some periods in different labs, where I met a lot of great scientists and good people with whom I shared my time. Thanks to Maite, Eelco, Lucas for taking care of me and to the nice people from their labs which gave me the possibility to feel home where I wasn't.

A really special thanks to the people, flatmates and friends, that took part in my "roman phase". Among them my gratitude especially goes to Carletta, for our infinitely long lunch breaks, Rossella and Stefano, for the plethora of crazy sketches I could use to write a book, to Iris, for her special enthusiastic way of living, to Fefi and Alessio, who shared with me labs, classes, and sometimes shopping times, to Roslen, for our doctoral gossip time, to Deborha, Ila, Luca, Daniela and Andrea, and to all those friends I forgot to mention.

My biggest debt of gratitude is for my family, mum and dad, for giving me and my sisters all they could, for not being stressful to me, for always giving me the freedom to choose whatever I wanted, to my sister Paola, you were, you are and you will always be a solid point of reference in my life, to my sister Elisa, for always taking care of me, no matter how distant we could be. A special thanks to my brother-in-law Fra, for supporting my sister all the time. A really special thanks goes to my nephew Samu, who still doesn't know how great is the joy he gave me while calling me "zia Lalla".

A tremendous thanks to my really big family, the multitude of uncles, cousins, relatives of every degree and type (don't get upset if I'm not mentioning all of you), which always were present in my life and always will be, I'm pretty sure. A special mention goes to my "cummarella" Nata, for her contagious laughter and her will to live, and my grandmothers, for being great women. Thanks to all of you for loving me the way I am.

I owe much also to Sandra e Bartolo, for welcoming me as a daughter and always make me feeling home.

Last but not least, Fabio, I don't have good words to sum up here my gratitude to you. Thanks for taking part of my life and my projects no matter how distant we could be, always making me feel the prettiest, the smartest, the best.

Lastly, I hope this PhD thesis will always remember me that I can count to some fortitude I not even hoped to have.

Ringraziamenti

Guardando indietro, posso quasi affermare con certezza che questa, per me, rappresenta la parte più complessa della stesura della tesi. Questo lavoro rappresenta la conclusione di un percorso di tre anni presso il reparto di Zoonosi Trasmesse di Alimenti dell'Istituto Superiore di Sanità, dove accanto ai momenti di sacrificio e impegni non sono mai mancati quelli di piacevoli distrazioni. Nella fase conclusiva di questo progetto, mi guardo indietro con gratitudine consapevole del fatto che non avrei potuto desiderare di meglio per la mia formazione, tanto scientifica, quanto umana.

Un grazie al prof. Paolo Visca, professore all'Università di Roma Tre e supervisore prima e durante il mio dottorato di Ricerca.

Un sentito ringraziamento va al dr. Alfredo Caprioli, per avermi innanzitutto ospitata nel suo Reparto presso l'Istituto Superiore di Sanità, per avermi incoraggiata nel mio percorso, e sempre ascoltata in caso di necessità, indipendentemente dalla sua mole di impegni.

Un grazie speciale va al mio supervisore dr. Stefano Morabito, per aver tentato di insegnarmi pazientemente come muovermi nell'affascinante mondo della Ricerca, per avermi sempre spronata a mettermi alla prova e a fare di meglio. Non ultimo lo ringrazio per il suo appassionato atteggiamento nel lavoro e nella vita, sempre divertente e positivo, e per aver spesso anche rappresentato un amico, oltre che un supervisore. Mi auguro solo di poterlo rendere orgoglioso del suo investimento..

Vorrei inoltre ringraziare di cuore le "droplets" Valeria, Rosy e Antonella. Valeria, per essersi presa cura di me nei miei primi passi in laboratorio, per aver rappresentato un solido e prezioso aiuto nei momenti di difficoltà, non solo lavorativi, per aver riempito di chiacchere, tanto scherzose quanto serie, il tempo trascorso insieme. Rosy, per essere stata una guida e un riferimento anzitutto, una consigliera, una preziosa amica, a seconda delle necessità. Antonella, per la sua freschezza e spontaneità, per aver spesso agito da "motivatore" per tutte noi, per il suo modo speciale di starmi vicina nei momenti difficili, e per le lacrime che verserà leggendo queste righe.

Un grande ringraziamento va a tutte le persone che hanno incrociato il mio percorso lavorativo spronandomi a dare e fare sempre meglio, dai numerosi ospiti stranieri e non che hanno frequentato il laboratorio, agli studenti, tirocinanti e quant'altro. Tra loro un grazie speciale va a Paola, per la sua sensibilità fuori dal comune, a Fernanda, per avermi fatta ridere come non mai, a Federica, per i suoi racconti esilaranti.

Devo molto della mia formazione umana e lavorativa alle mie esperienze all'estero e alle numerose persone che ne hanno fatto parte, dai supervisori Maite, Eelco, Lucas e alle persone che quotidianamente hanno incrociato e condiviso il mio cammino: grazie per avermi ospitata, seguita e spronata, e spesso fatta sentire a casa quando non lo ero.

Un grazie speciale alle persone che fanno parte o hanno fatto parte della mia fase "romana", i miei coinquilini e amici. Tra loro un grazie speciale a Carletta, per i nostri pranzi infiniti, a Rossella e Stefano, per avermi fornito tale quantità di spunti di ordinaria follia da poter scrivere un libro, a Iris, per il suo influsso sempre positivo, a Fefi e Alessio, che tra una lezione, un laboratorio, una spedizione di shopping, non sono mai mancati, a Roslen, per i nostri sfoghi sulle attività dottorali e non solo, a Deborha, Ila, Luca, Daniela e Andrea, e tutti gli altri che non sono in questa lista..

Un sentito grazie va poi alla mia famiglia, a mia madre e mio padre anzitutto, per aver dato tutto quello che avevano per le loro figlie, per non essere mai stati pressanti, per avermi lasciato sempre la libertà di scegliere quello che volevo, a

mia sorella Paola, per aver rappresentato sempre e comunque un punto di riferimento nella mia vita, a mia sorella Elisa, per essersi sempre presa cura di me, indipendentemente dalla distanza. Un grazie inoltre va a mio cognato Francesco, che sopporta mia sorella tutti i giorni…Il mio grazie più sentito va però a mio nipote Samuele, che nel chiamarmi "zia Lalla" non sa ancora di avermi procurato una delle gioie più grandi della terra. Un enorme grazie va poi alla mia enorme famiglia allargata, la carrellata di zii, cugini, parenti di ogni grado e categoria (non me ne vogliano se non li nomino tutti), che sono sempre stati presenti nella mia vita e che sono sicura, sempre ne faranno parte. Una menzione speciale merita la mia "cummarella" Nata, per la sua risata e la sua voglia di vivere contagiose, e le mie nonne, per il loro esempio di donne fuori dal comune. Grazie a tutti per avermi sempre voluta bene per quella che sono.

Devo molto a Sandra e Bartolo, per avermi accolta come una figlia e fatto sempre sentire a casa.

Non ultimo, Fabio, non so come riassumere qui in due righe la gratitudine nei tuoi confronti. Grazie per aver preso parte alla mia vita e ai miei progetti indipendentemente da quanto fossimo distanti, facendomi sempre sentire la più bella, la più intelligente, la più capace.

E infine, un monito a me stessa, spero che questa tesi possa sempre ricordarmi che posso contare su una forza d'animo che non credevo nemmeno di avere.

Table of Contents

CHAPTER 1

Introduction

1.1 Escherichia coli

Escherichia coli is a Gram-negative facultative anaerobe bacterium that colonizes the intestine of mammals during the first months of life becoming an important member of the gut microflora (Tenaillon *et al.*, 2010). The relationship between *E. coli* and its host is based on mutual benefits, in which the former obtains a steady supply of nutrients, protection from stresses and the possibility to be disseminated in the environment and to other hosts with faeces (Tenaillon *et al.*, 2010), while favouring the turnover of intestinal epithelium in the latter, promoting an healthy immune response against pathogens.
E. coli has a biphasic lifestyle, being able to persist for prolonged periods outside the host in environmental niches, such as water and soil (Vogeleer *et al.*, 2014).
When living within the human body, *E. coli* is generally confined to the intestinal lumen; however in the immune-suppressed host and/or when intestinal barriers are injured, *E. coli* strains can reach other body compartments causing disease (Nataro *et al.*, 1998a). Additionally, some strains evolved the ability to cause disease in healthy individuals in their own right, ranging from mild, self-limiting illness to life-threatening forms of infection (Nataro *et al.*, 2001).
Pathogenic *E. coli* cause three general syndromes including urinary tract infections, sepsis/meningitis and diarrheal diseases and are divided into two main categories: Extraintestinal Pathogenic *E. coli* (ExPEC) and Diarrheagenic *E. coli* (DEC). More in detail, strains presenting similar mechanisms of pathogenesis or causing similar clinical symptoms have been collected into sub-groups called pathotypes (Donnenberg *et al.*, 2001). Intuitively, ExPEC are responsible of infections occurring outside the gastrointestinal tract. These include the genito-urinary tract, the central nervous system, the circulatory and the respiratory systems (Russo *et al.*, 2003). ExPEC pathotypes include uropathogenic *E. coli* (UPEC), strains associated with bacteraemia and sepsis (SePEC) and neonatal meningitis-associated *E. coli* (NMEC). UPEC are responsible for 70–95% of community-acquired urinary tract infections in the general population in the USA and approximately 50% of nosocomial infections, hence accounting for substantial morbidity, and burden of disease (Foxman, 2010). NMEC cause 20-40% of neonatal meningitis occurring in the USA with a mortality rate of approx. 8% (Smith *et al.*, 2007).
The DEC group includes strains causing gastroenteritis in humans and animals and are responsible for a range of diseases causing mild to moderate symptoms, including watery and protracted diarrhoea, to more severe forms such the haemorrhagic colitis and the life-threatening haemolytic uremic syndrome (Tozzoli *et al.*, 2014b). Being the main focus of this thesis, the DEC group and the related pathotypes will be treated more in detail in the next sections.

1.2 Diarrheagenic *Escherichia coli* pathotypes

The currently recognized DEC pathotypes are: enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC), diffusely adherent *E. coli* (DAEC), enteroaggregative *E. coli* (EAggEC), enteropathogenic *E. coli* (EPEC), and Shiga toxin-producing *E. coli* (STEC), also referred to as Verocytotoxin-producing *E. coli* (VTEC).

1.2.1 Enterotoxigenic *Escherichia coli*

ETEC strains are a prevalent cause of diarrhoea among children in low-income countries and an important cause of the traveller's diarrhoea, a disease affecting people from industrialized countries travelling to developing regions (Northey *et al.*, 2007). In addition, infections caused by these microorganisms contribute substantially to delayed childhood growth and malnutrition in developing countries (Qadri *et al.*, 2007). ETEC elaborate at least one member of two defined groups of enterotoxins: STs (heat-stable toxins) and LTs (heat-labile toxins), both causing an increase of intracellular level of the second messengers' cyclic nucleotides and an unbalanced movement of ions in small intestine epithelial cells, which determines the secretory watery diarrhoea (Kaper *et al.*, 2004). Besides toxins production, ETEC pathogenicity relies upon the presence of virulence factors such as adhesins, generically named as colonization factors (CFs or

CFAs). More than twenty CFs variants have been described until now, displaying different types of structures (fimbriae, fibres and non-fimbrial) (Del Canto *et al.*, 2014). To date there is no effective prophylaxis for ETEC infections, even though CFs seem to be suitable antigens and are under investigation for vaccines formulations (Gaastra *et al.*, 1996).
ETEC strains require a high infectious dose to cause infections, thus making direct person-to-person transmission rare. On the other hand, poor drinking water quality, lack of a sewage system and feeding of supplementary foods are risk factors for ETEC infection. Additionally, contaminated food, mainly raw vegetables and soft cheeses, have been reported as common sources of sporadic cases and outbreaks in industrialized countries (Tozzoli *et al.*, 2014b).

1.2.2 Enteroinvasive *Escherichia coli*

EIEC are biochemically, genetically and pathogenically closely related to *Shigella* spp. from which are distinguished by a few minor biochemical tests (Kaper *et al.*, 2004). EIEC may cause an inflammatory colitis and dysentery. However, in most cases the strains belonging to this pathotype elicit a watery diarrhoea often indistinguishable from that due to infection by other pathogenic *E. coli* (Nataro *et al.*, 1998a). EIEC strains present a peculiar mode of pathogenesis characterised by the vacuole-mediated penetration of epithelial cells of the distal large bowel followed by the lysis of the endocytic vacuole, intracellular multiplication and movement into adjacent cells (Sansonetti, 2002). The ability of EIEC to invade and destroy the intestinal tissue is largely associated with the presence of a plasmid (pINV), also present in *Shigella* spp., which harbours virulence factors such as the invasion antigens named as IpaA to IpaH (Lan *et al.*, 2004) and a type III secretion system (T3SS), responsible for the translocation of multiple proteins into the eukaryotic cells which trigger signalling events, cytoskeleton rearrangements and lysis of the endocytic vacuole (Tran Van Nhieu *et al.*, 2000).
EIEC are human pathogens and infection occurs *via* the oral-faecal route following a person-to-person transmission (Tozzoli *et al.*, 2014b).

1.2.3 Diffusely Adherent *Escherichia coli*

DAEC are defined by the presence of a diffuse pattern of adhesion to HEp-2 cells. These strains induce a cytopathic effect characterised by the development of long cellular extensions, which wrap around the adherent bacteria. This requires the production of proteins belonging to the Dr family of adhesins, termed Dr Adhesion factor (DAF) (Kaper *et al.*, 2004). DAEC strains have been implicated as a cause of diarrhoea particularly in children elder than 12 months (Scaletsky *et al.*, 2002). However, their precise role in diarrheal diseases, their classification and epidemiology of the infections are still unclear (Tozzoli *et al.*, 2014b).

1.2.4 Enteroaggregative *Escherichia coli*

EAggEC strains can cause watery, mucoid, secretory and persistent diarrhoea with low-grade fever and little or no vomiting in children and adults of both developing and industrialized countries (Bhan *et al.*, 1989). Such strains are able to grow and survive embedded in a mucus-containing biofilm, thus strongly adhering to the human intestines (Nataro *et al.*, 1998b). The striking feature of these strains is the "stacked-brick" adhesion pattern to epithelial cells, in which the bacteria both bind to the epithelial cells' surface and aggregate one another due to the production of a wide range of fimbriae (Harrington *et al.*, 2006). These include the aggregative adhesion fimbriae (AAF/I to AAF/V), whose genes are encoded on 55-65 MDa plasmids (pAAs). AAF genes expression, as well as that of other virulence factors, is regulated by the master and plasmid-conveyed transcriptional activator AggR (Dudley *et al.*, 2006). The same plasmid carries other virulence factors such as the gene coding for the antiaggregation protein, or dispersin (Aap), able to neutralize the LPS negative charges and thus favouring the adhesion, and the *aat* operon encoding the ABC transporter responsible for dispersin secretion. Beside the colonization and the biofilm production,

EAggEC pathogenicity relies on the secretion of toxins, which play an important role in causing the watery diarrhoea (Harrington *et al*., 2006). These include the serine protease autotransporters of the Enterobateriaceae (SPATEs), the EAST-1 cytotoxin, ShET1 (Hebbelstrup Jensen *et al.*, 2014). EAggEC reservoir has not been determined, but it is generally accepted to be human (Huppertz *et al.*, 1997). Infections are transmitted *via* the oral-faecal route through inter-human contacts or the ingestion of contaminated water or food (Jiang *et al.*, 2002).

1.2.5 Enteropathogenic *Escherichia coli*

EPEC, the first DEC pathotype ever described, is the main causative agent of infantile diarrhoea in developing countries (Nataro *et al.*, 1998a). An EPEC strain is defined by its ability to cause the "attaching and effacing" (A/E) lesion to the epithelial cells (Kaper, 1996). Two sub-groups can be individuated: typical EPEC (tEPEC) and atypical EPEC (aEPEC), genetically distinguished by the presence, in the former, of a large plasmid called EAF (EPEC Adherence Factor) and in the majority of the latter, of a large virulence plasmid resembling that present in some strains belonging to another DEC pathotype, the Verocytotoxin-producing *E. coli* (VTEC; see below). The two sub-groups also present differences in the epidemiology of the infections, with aEPEC being characterised by transmission routes more similar to those of the infections caused by VTEC, which are typically foodborne zoonoses, and the tEPEC infections being mainly spread through inter-human contacts and the usual oral-faecal routes. Both EPEC groups may cause persistent chronic diarrhoea, (Donnenberg, 1995) with more severe acute forms sometimes reported (Bower *et al.*, 1989). Vomiting and low-grade fever are also common symptoms of EPEC infection (Boisen *et al.*, 2014).

The histopathology of EPEC infection includes the formation of a peculiar lesion to the enterocyte, termed the attaching and effacing (A/E), which results from a multi-steps mechanism. The bacteria first interact with the enterocyte layer by means of a bundle-forming pilus (BFP), whose gene is encoded on the EAF plasmid. In the second step an adhesin, the intimin, is produced and exposed on the bacterial surface, while several effectors are released into the host cell *via* a T3SS, determining the disruption of microvilli, accumulation of actin and the formation of cup-like pedestals upon which the bacteria sit (Fig. 1.1). In the third step, the intimate attachment is mediated by the strong interaction between the intimin and Tir, the intimin receptor produced by the bacterium, translocated through the T3SS and eventually exposed on the host cell surface. The ability to cause this lesion is conferred by the presence of 41 genes conveyed by a 35-40 Kb pathogenicity island, called the Locus for Enterocyte Effacement (LEE) (Donnenberg *et al.*, 2001, Kaper *et al.*, 2004).

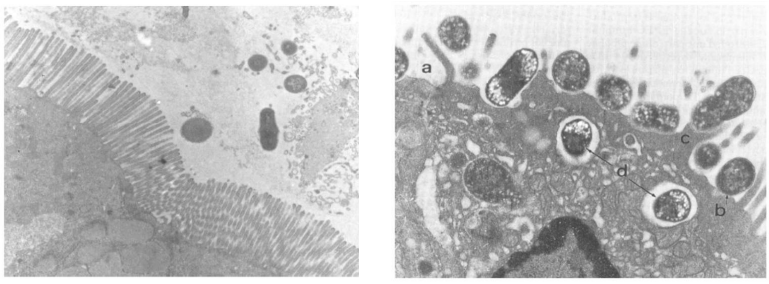


Fig. 1.1. Attaching and effacing lesions produced by an enteropathogenic E. coli (EPEC) in the ligated loop intestinal assay in rabbit. Healthy enterocytes with microvillus layer (left); A/E lesion (right) from: a) few residual microvilli; b) intimately adherent E. coli; c) pedestal formation beneath the adherent bacteria; d) sometimes internalization of the bacteria into the enterocytes (Piérard et al., 2012).

Additional virulence factors of both tEPEC and aEPEC have been described outside the LEE locus such as the large protein lymphostatin (LifA), also known as Efa1, that inhibits lymphocyte activation (Klapproth *et al.*, 2000) and enhances the bacterial adhesion (Tatsuno *et al.*, 2000). In the tEPEC, the large plasmid EAF encodes, besides the BFP, the *per* locus (plasmid-encoded regulator), whose genes are involved in the regulation of several virulence factors.

1.2.6 Verocytotoxin-producing *Escherichia coli*

VTEC were first recognized as cause of human infections in the early '80s during two outbreaks of bloody diarrhoea occurred in the USA and associated to the consumption of undercooked hamburgers at a fast-food chain (Riley *et al.*, 1983). Stool cultures from patients yielded *E. coli* isolates belonging to a rare serotype, O157:H7. The disease was designated haemorrhagic colitis (HC) and the causative agents Enterohaemorrhagic *E. coli* (EHEC).

VTEC are zoonotic pathogens with cattle being recognized as the major reservoir (Caprioli *et al.*, 2005). Transmission of VTEC infections occurs through the consumption of contaminated food and water, but person-to-person transmission or by direct contact with animals or animal manure have also been reported, although rarely (Heuvelink *et al.*, 2002).

Nowadays, a wide variety of vehicles for VTEC infections has been described, including hamburger and other meat products, water and unpasteurised milk, cantaloupe melon, apple juice, leafy vegetables and several types of sprouts among others (Kaper *et al.*, 2004). VTEC cause large community outbreaks, sometimes with severe outcomes, facilitated by a very low infectious dose estimated to be around ten colony-forming units (CFU) for VTEC O157. In 1996 in Japan about 10,000 people became infected with VTEC O157:H7 following the consumption of radish sprouts (Watanabe *et al.*, 1999). In 2000, in Canada 2,300 people acquired VTEC O157:H7 through the consumption of contaminated water and seven died (Holme, 2003). In 2011, in Germany and France, more than 4,000 people were infected by a VTEC O104:H4 strain, 900 developed HUS, 54 died (Karch *et al.*, 2012).

The main virulence feature of VTEC strains is the ability to elaborate potent cytotoxins known as Verocytotoxins (VTs). These are AB_5 toxins, consisting of five identical B subunits responsible for the binding to the cellular receptor and a single A subunit, which cleaves the rRNA of target cells, causing cell death by blocking protein synthesis (Melton-Celsa *et al.*, 1998). VTs are also called Shiga-Toxins (Stxs) because of their similarity to the Shiga-toxin produced by *Shigella dysenteriae* type I, and VTEC are therefore also termed STEC (Shiga-toxin producing *E. coli*). There are two main antigenic types of VTs, VT1 and VT2, including three subtypes of VT1 (VT1a, VT1c and VT1d) and seven subtypes of VT2 (VT2a, VT2b, VT2c, VT2d, VT2e, VT2f and VT2g) (Scheutz *et al.*, 2012). The VT-coding genes (*vtx*) are conveyed by lambdoid bacteriophages integrated in the bacterial chromosome (Ogura *et al.*, 2007), which are usually maintained in a lysogenic state but retain the capability to enter the lytic cycle, multiply and move to other hosts spreading the *vtx* genes (Plunkett *et al.*, 1999). Each VT-phage carries one *vt*x operon consisting of the two subunit genes, *vt*x*A* and *vt*x*B*, which encode the complete AB_5 holotoxin. VTEC can produce either VT1 or VT2 alone or both and in different subtypes combination (Muniesa *et al.*, 2014). VTs are produced in the colon and may reach the kidney's glomerular microvasculature *via* the bloodstream, where through direct apoptotic activity and pro-inflammatory actions damage the endothelial cells causing occlusion. This can lead to the Haemolytic Uremic Syndrome (HUS), whose main symptoms are haemolytic anaemia, thrombocytopenia and potentially fatal acute renal failure (Andreoli *et al.*, 2002). The estimated rate of fatality due to HUS is 3-5% (Karch *et al.*, 2005). Other sequelae of VTEC infection may include neurological consequences due to the action of VTs on the endothelium of the brain vessels.

The treatment of VTEC infections is mainly supportive and consists in rehydration and in the dialysis for the treatment of HUS. In fact, administration of antibiotics is not recommended, since it has been proposed

to cause the progression of infection towards the most severe forms, probably due to the antibiotic-dependent enhanced VTs production and its massive release into the bloodstream (Kimmitt *et al.*, 2000). The presence of VT-bacteriophages seems not to be sufficient for VTEC to cause disease; VTEC strains isolated from the most severe forms of infection, HC and HUS, also carry additional virulence factors involved in the colonization, and more in general, in the pathogenetic mechanism (Tozzoli *et al.*, 2014b). Most of VTEC strains associated with HUS and with epidemic diseases possess the LEE locus, a genetic element shared with EPEC and responsible for the induction of the histopathological lesion called "attaching and effacing" (A/E) (McDaniel *et al.*, 1995, Karmali *et al.*, 2003).

Although serotype O157:H7 has been implicated in the largest VTEC outbreaks (Karmali, 1989, Griffin *et al.*, 1991), there is growing concern about the risk posed to human health by non-O157 VTEC serotypes, more than 400 of which have been so far associated with human illness (Tozzoli *et al.*, 2014b).

LEE-negative VTEC strains, in some cases, have also been associated with serious human disease (Nataro *et al.*, 1998a, Johnson *et al.*, 2006, Käppeli *et al.*, 2011). Such strains usually possess alternative virulence-associated genes, such as the adhesin Saa, encoded on the plasmid pO113 described in LEE-negative VTEC of serogroup O113 (Paton *et al.*, 2001). In addition to *saa*, the genes *sab*, *epeA* and *subAB*, encoding respectively a SPATE exhibiting protease and mucinase activity, a protein contributing to adherence and biofilm formation, and the Subtilase cytotoxin, are carried by the same plasmid (Paton *et al.*, 1999, Steyert *et al.*, 2012).

The continuously growing use of genome sequencing technologies paved the way to a more comprehensive knowledge of the VTEC genomes. The determination of the complete genome sequence of *E. coli* O157:H7 strains EDL 933 and Sakai at the beginning of 21st century, showed that in these VTEC strains, the chromosome contains more than 170 genomic islands which are not present in the *E. coli* K-12 MG1655 laboratory strain genome and that 33% of them harbour genes with unknown functions (Hayashi *et al.*, 2001, Perna *et al.*, 2001). Among these genomic islands, a 22-kb PAI designated OI-122 in VTEC O157:H7 EDL933 strain and SplE3 in the Sakai strain, carries the 5' of the *efa1/lifA* gene (Morabito *et al.*, 2003), which is involved in the repression of host lymphocyte activation (Klapproth *et al.*, 2000). PAI OI-122 is strongly associated with the LEE locus in both VTEC and EPEC strains, and in many of them is physically linked to the LEE locus itself in a mosaic PAI (Morabito *et al.*, 2003). It has been proposed that the LEE locus and PAI OI-122 may have been acquired as a unique larger PAI and that they separated later on in some strains, following genetic rearrangement events (Morabito *et al.*, 2003).

Most virulent VTEC strains also possess a large virulence plasmid, called pO157 in VTEC O157 strains, conveying the genes for the production of the enterohaemolysin, a toxin favouring the release of haemoglobin from red blood cells during infection, thus providing a source of iron for the bacteria (Schmidt *et al.*, 1995) and other putative virulence genes such as *espP* and *katP* encoding a serine protease and a catalase-perossidase, respectively (Brunder *et al.*, 2006). The pO157 also hosts the gene *toxB* (Tozzoli *et al.*, 2005, Michelacci *et al.*, 2014), whose product proved to contribute to VTEC O157 adherence to Caco-2 cultured cells by promoting the production and/or the secretion of type III secreted proteins (Stevens *et al.*, 2004).

In the attempt to come to a classification of the various VTEC types, the VTEC serotypes have been distributed into five categories termed "seropathotypes" (SPTs) and indicated with the letters from A to E, in a descending order of pathogenicity (Karmali *et al.*, 2003). Such a distinction has been based on their reported frequency in the human illness, the association with severe disease and outbreaks, and the presence of mobile genetic elements conferring virulence genes such as the LEE locus and the PAI OI-122 (Karmali *et al.*, 2003).

Studies aiming at comparing the different SPTs for the identification of the whole genomic asset conferring to VTEC the full pathogenicity, have led to the description of another PAI, PAI OI-57, which seems to be

associated with the most virulent SPTs, A and B (Imamovic *et al.*, 2010). This PAI harbours two genes whose products are annotated with a putative function: *adfO*, coding for a factor described to increase the ability of VTEC O157 to adhere to cultured HeLa cells (Ho *et al.*, 2008), and *cfk*, a phage-associated bacterial cell killing factor (Perna *et al.*, 2001).

1.2.7 Enteroaggregative-Haemorrhagic *Escherichia coli*

Although not yet recognized as an official *E. coli* pathotype, the Enteroaggregative Haemorrhagic *Escherichia coli* group (EAHEC) deserves to be mentioned. Such an *E. coli* subpopulation showcases virulence features transversal to two different DEC types: VTEC and EAggEC. The proposal for such a pathotype was ratified in 2011 after a large outbreak of haemorrhagic colitis and HUS that plagued Germany and, to a lesser extent France, with more than 4,000 cases of infection, 900 HUS, and 54 deaths (Frank *et al.*, 2011). The outbreak was characterised by an unusual high rate of infections progressing to HUS (22% in spite of the usual 5-10%) and the infectious agent belonged to a before rarely reported *E. coli* serotype, O104:H4 (Frank *et al.*, 2011). The outbreak strain carried the VT2a-phage and induced the "stacked-brick" pattern of adhesion to cultured Hep-2 cells, typical of EAggEC strains (Scheutz *et al.*, 2011). In fact, it carried genes peculiar of the EAggEC pathotype, such as the aggregative adherence fimbriae type 1 and several SPATEs (Steiner, 2014).

After the German outbreak in 2011, the scientific community looked retrospectively at the reported HUS cases linked to infections with atypical VTEC types or browsed the scientific literature with the aim to assess if other EAHEC cases of infection could be retrieved. It turned out that the first ever reported EAHEC strain caused eight HUS cases in France in 1992 and belonged to O111:H10 serotype (Morabito *et al.*, 1998). Additionally, in the time-span 1992-2012, at least six sporadic cases of EAHEC-associated HUS were observed as caused by EAHEC strains belonging to three different serotypes such as O86: HNM, O104:H4 and O111:H21 (Morabito *et al.*, 1998, Iyoda *et al.*, 2000, Sca*via et al.*, 2011, Dallman *et al.*, 2012). Furthermore, an O15 VT2-producing strain positive for the presence of Enteroaggregative markers has been described in a patient with septicaemia (Wester *et al.*, 2013). Finally, an EAHEC strain O127:H4 strain has been recently isolated from four HUS cases occurred in northern Italy in 2013 (Tozzoli *et al.*, 2014a).

The appearance of the first EAHEC raised the question if it was indeed a new pathogenic group of *E. coli* or we were rather observing accidental sporadic events of VT-bacteriophages acquisition by classical EAggEC strains.

The experience of the German outbreak and the following reported existence of at least six different EAHEC serotypes, as well as the growing number of reports of *E. coli* other than VTEC producing VTs, such as the ExPEC strains isolated from patients with bacteraemia (Wester *et al.*, 2013) or ETEC strains displaying the presence of the *vtx*-coding genes (Tozzoli *et al.*, 2014a), seem to indicate that *E. coli* pathotypes should not be considered as rigidly separate entities but rather *E. coli* sub-populations dynamically exchanging part of their genomic traits causing the continuous emergence of hybrid groups, sometimes with augmented pathogenicity, as in the cases of EHEC and EAHEC (Fig. 1.2).

It´s now generally accepted that *E. coli* genome as a whole, also termed pangenome, is constituted of a ¨backbone¨ of around 2,000 genes shared by all *E. coli* and a supplement of accessory genes, generally carried by Mobile Genetic Elements (MGE), peculiar to each subpopulation (Franz *et al.*, 2014).

MGEs strongly contribute to the high flexibility and dynamicity of *E. coli* genome and constituted the engine that powered the evolution of the different DEC pathotypes, therefore will be treated more in detail in following sections.

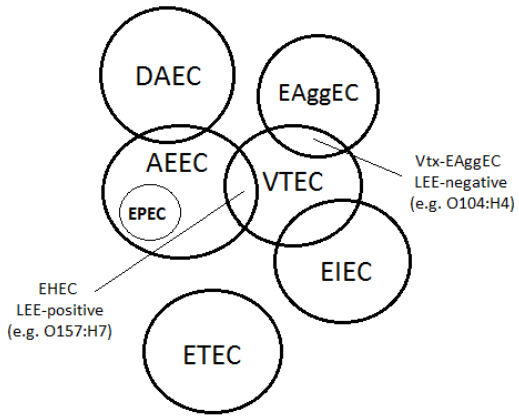


*Fig. 1.2. Relationships between human DEC pathotypes. DAEC: Diffusely-adherent E. coli; VTEC: Verocytotoxin-producing E. coli; AEEC: Attaching and effacing E. coli; EIEC: Enteroinvasive E. coli; EPEC: Enteropathogenic E. coli; EHEC: Enterohaemorrhagic E. coli; ETEC: Enterotoxigenic E. coli; EAEC: Enteroaggregative E. coli; Vtx-*EAggEC*: Shiga toxin-producing Enteroaggregative E. coli. Areas of overlapping between circles represent the fractions of the genome shared by the related pathotypes (adapted from Franz et al., 2014).*

1.3 Epidemiology of diarrheagenic *Escherichia coli* infections

VTEC represent the only pathogenic group of *E. coli* that has a definite zoonotic origin with cattle being recognized as the major reservoir (Caprioli *et al.*, 2005). Even though VTEC can be found in the gut of numerous animal species, ruminants are considered as the major animal source of strains highly virulent to humans (Caprioli *et al.*, 2005).

Transmission of VTEC infections to man occurs through the consumption of food and water, including in the latter the exposure linked to recreational activities. Food of animal origin can be primarily contaminated during the transformation process, e.g. the carcasses and meat at the slaughterhouses or milk during milking procedure. Vegetables can become contaminated following an environmental pathway including the use of ruminants' manure to fertilize fields where crops are grown. Although most of VTEC infections, especially those caused by O157 strains have been linked to exposure to a food vehicle or water, person-to-person transmission or by direct contact with animals or animal manure have also been reported (Heuvelink *et al.*, 2002).

VTEC infections are not mandatorily notified in most of the countries worldwide, resulting in a massive underestimation. In the European Union, the cases of infections caused by VTEC are communicated, voluntarily by the EU Member States, to the European Centre for Disease Prevention and Control (ECDC), which collects the information and elaborate trends and incidence data. In the last published Annual Epidemiological Report, referring to the data of 2011 year, 9,534 VTEC infections have been reported from 27 European countries (ECDC, 2014). This number is 2.5 higher than the previous year (Fig 1.3) due to the occurrence of the large German outbreak of VTEC O104:H4 infections (See section 1.2.7). A total of 1,006 of the reported cases were HUS, showing an incidence of 11% for this severe form of the infection. Twenty-eight per cent of HUS cases were reported in 0-4 years old children with O157 and O26 as prevalent serogroups, followed by 25-44 years old adults with O104 being the predominant serogroup (ECDC, 2014).

As for the other diarrheagenic *E. coli*, ETEC, EAggEC, EPEC and EIEC, given their inter-human circulation and the oral-faecal route of transmission of the infections (See section 1.2), their circulation is

more common in developing countries where the high incidence of the disease caused by these pathotypes of *E. coli*, in conjunction with the poor hygienic conditions and the lack of effective human sewage treatments is responsible for their environmental diffusion and persistence. ETEC, EPEC, EAggEC are the most important enteric pathogens in children below five years of age and are responsible for roughly a half of all the diarrheal episodes in developing countries (Clarke, 2001).

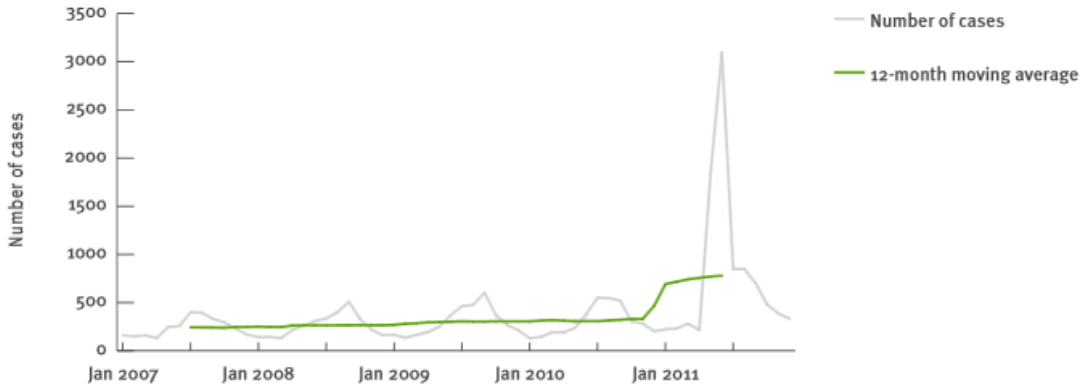


Fig.1.3 Trend and number of VTEC cases reported in the EU, 2007-2011 (ECDC, 2013).

Dutta and colleagues (Dutta *et al.*, 2013) reported the results of an investigation on 3,826 stool specimens collected from acute diarrheal patients hospitalized at the Infectious Diseases Hospital in Kolkata, India, and found that the major prevalence of infections was due to EAggEC (5.7%), followed by ETEC (4.2%) and EPEC (1.8%). Other authors reported variable rates of EPEC isolation ranging from 12.6 to 44.9% in children under five years with diarrhoea in Iran (Jafari *et al.*, 2012). In the same country, the same study reported ETEC detection rates from less than 10% up to 33% (Jafari *et al.*, 2012).
A recent case-control study described a prevalence of EAggEC of 41% in children under five years in north-eastern Brazil compared to previous data ranging from 2 to 12% (Lima *et al.*, 2013). In Guinea-Bissau the most common bacteria isolated from children under two years with persistent diarrhoea was EAggEC (Valentiner-Branth *et al.*, 2003).

Among the human-borne *E. coli* pathotypes, the EAggEC show the more variable epidemiology. This pathotype has been reported as a leading cause of travellers' diarrhoea, a common cause of persistent diarrhoea among HIV-infected people as well as of endemic paediatric diarrhoea in industrialized and developing countries (Boisen *et al.*, 2014). In a large study in England, including more than 3,600 cases of diarrhoea, EAggEC were the second most common bacterial cause of gastroenteritis (Tompkins *et al.*, 1999).

1.4 Mobile Genetic Elements and their role in bacterial evolution

E. coli genome vary in size from 4.7 Mb to 6.5 Mb (Hazen *et al.*, 2012). The studies on the *E. coli* pangenome indicate that this species is continuously evolving through horizontal gene acquisition and diversification (Dobrindt *et al.*, 2010). Changes occur also through point mutations, deletions and single nucleotide variations (Ochman *et al.*, 2000). These events contribute to the genome optimization (Fig. 1.4) and eventually lead to the adaptation of the microbe, which gains new niches, escape from host immune defences and acquire new resistances to antimicrobials.
The genome plasticity and the resulting variability have made a significant contribution to the successful emergence of new pathogenic *E. coli* (Leopold et al., 2014).

A striking feature of pathogenic *E. coli* is the presence of genes that encode virulence factors on mobile genetic elements (MGEs) (Kaper *et al.*, 2004). MGEs include plasmids, genomic islands (GIs), pathogenicity islands (PAIs), bacteriophages, integrons, transposons and insertion sequence elements (ISs) (Dobrindt *et al.*, 2010).

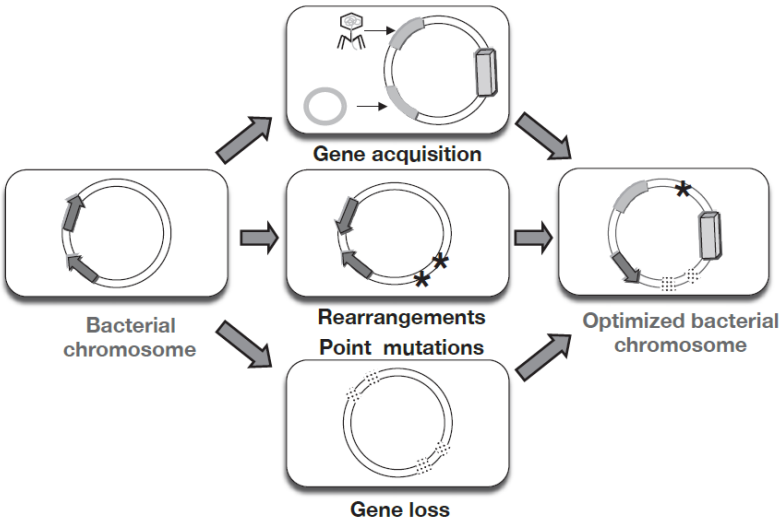


Fig.1.4 Mechanisms of genome optimization (Leopold et al., 2014).

1.4.1 Genomic and Pathogenicity Islands

PAIs and GIs form a distinct super-family of MGEs (Hacker *et al.*, 1997) that integrate into and excide from the host chromosome making the occurrence of horizontal gene transfer possible.
GIs typically integrate in the host chromosome in the proximity of tRNA genes, possess a different G+C content compared to the average of the genome backbone, and often carry cryptic or functional genes encoding factors related to the self-mobilization, such as integrases and transposases (Kaper *et al.*, 2004). Different types of GIs have been described so far, some convey genes involved in the metabolism and enhancing the microbial fitness, other confer drug resistance or are involved in the pathogenetic mechanism. The latter are termed Pathogenicity Islands (PAIs). PAIs encode a wide variety of virulence factors, such as toxins, adhesins, iron uptake systems, secretion systems and strategies to escape from the host immune defence mechanisms (Dobrindt *et al.*, 2010).
Some PAIs contain regions showing homology to other MGEs, such as bacteriophages and plasmids, suggesting that they may have derived from events of integration of such MGE in the host chromosome leading to a more stable association and vertical replication (Dobrindt *et al.*, 2004).
The presence in the GIs of genes displaying a certain homology in the DNA sequence with other ORFs facilitates the genetic re-arrangement events between different islands and within the same island, thus boosting the plasticity of *E. coli* genome.

1.4.2 Bacteriophages with emphasis on VT-phages

MGEs are the engine of bacterial diversity and bacteriophages in particular seem to have the deepest impact on the evolution of pathogenic bacteria (Campbell, 1996). As a matter of fact, the spread of virulence-associated genes by lysogenic phages is a well-known phenomenon in both Gram-negative and Gram-positive bacteria (Boyd *et al.*, 2002). The availability of complete *E. coli* genome sequences shows a high diversity in the amount of prophage-related sequences within this species (Dobrindt *et al.*, 2010). The modular structure of bacteriophages genome has been proposed since the early studies on this topic (Susskind *et al.*, 1978). According to this theory, bacteriophages are constituted by several functional

modules that can be interchanged between each other either by homologous recombination at specific sequences or by non-homologous recombination at random points (Muniesa *et al.*, 2014).
Among the several phages and phage-related sequences in *E. coli* genome, the Verocytotoxin encoding-bacteriophages (VT-phages) played, and may still be playing, a pivotal role in the evolution of the different *E. coli* pathotypes. In fact, besides VTEC, the ability of producing VT seems to be shared with almost all the pathogenic groups of strains belonging to this species (Scavia *et al.*, 2011, Wester *et al.*, 2013, Tozzoli *et al.*, 2014a). VT-phages constitute a heterogeneous group of temperate lambdoid phages that harbour *vtx* genes. VT-phages described so far belong to the Myoviridae and Siphoviridae virus families (Schmidt, 2001), vary in size genomes and composition, morphology, host range, integrate in different host chromosomal loci and differentially influence the host pathogenicity (Muniesa *et al.*, 2014, Tozzoli *et al.*, 2014a).
VT-phages alternate a lysogenic state, in which the phage genome is silently integrated into the host chromosome, with a lytic cycle, characterised by replication and phage particle formation, followed by the lysis of the host cell. VT-phages are then spread to other host cells.
VT-phages genomes have been characterised by restriction analysis and mapping experiments since the early '80s and lately by whole genome sequencing. Reported genome sizes range from 42 Kb to more than 60 Kb in phages isolated from different VTEC strains (Schmidt, 2001, Muniesa *et al.*, 2014).
The number of available VT-phages genome sequences is growing and several comparative genomics studies have been published with the aim of highlighting similarities and differences between them and understanding the basis and the factors influencing their heterogeneous biology (Fig. 1.5) (Ahmed *et al.*, 2012, Laing *et al.*, 2012, Smith *et al.*, 2012, Steyert *et al.*, 2012, Grande *et al.*, 2014). The genetic structure of VT-phages mainly corresponds to that of phage lambda, with an early region comprising genes involved in the regulation of lysogenic-lytic cycle and a late region containing all the genes necessary for packaging and phage particle formation. Vtx-coding genes are located in the late region downstream the anti-terminator site of Q gene and upstream the lysis cassette. The position of *vtx* genes is conserved and their transcription is boosted when lytic cycle is induced.

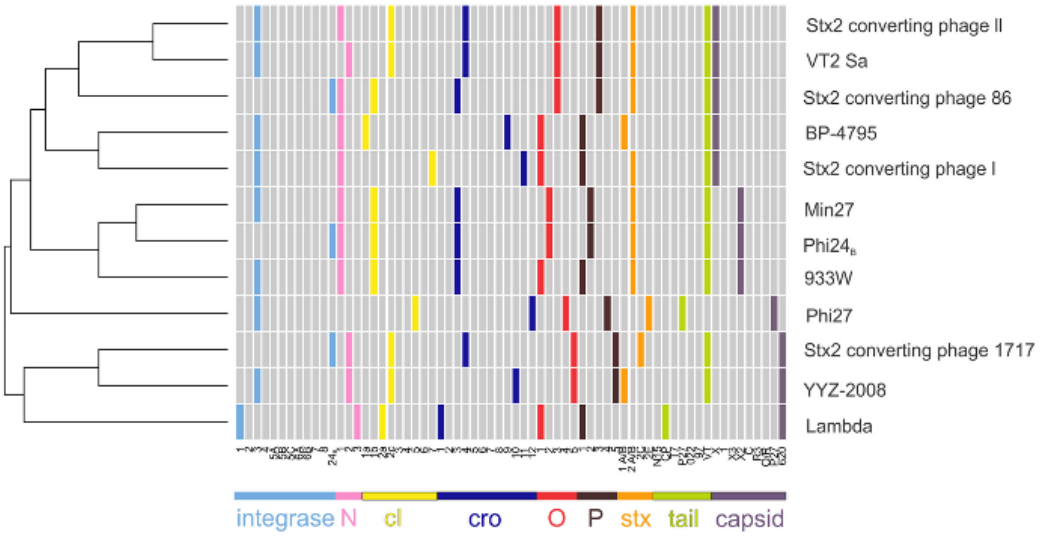


Fig. 1.5 Multi-genome comparison of VT-phages. Variants of the same gene are indicated with the same colour (Smith et al., 2012).

Lytic cycle can be induced by exposure to stress inducing agents, such as many classes of antibiotics, UV light, high hydrostatic pressure, H_2O_2 or occur spontaneously following SOS-independent mechanisms (Imamovic *et al.*, 2012).
VT-phages have been described in a number of Enterobacteriaceae including, beside different pathotypes of *E. coli* (Karch *et al.*, 1999, Tozzoli *et al.*, 2014a), *Shigella dysenteriae type I*, *Shigella sonnei, Shigella flexnerii* (Strauch *et al.*, 2001, Gray *et al.*, 2014), *Citrobacter freundii* (Schmidt *et al.*, 1993, Tschape *et al.*, 1995) and *Enterobacter cloacae* (Paton *et al.*, 1996).
Free VT-phages can be found in the environment, including water and sewage systems, in which they can survive for prolonged periods (Muniesa *et al.*, 2014). Their presence in the environment increases the risk of the emergence of bacterial pathogens, especially new pathogenic *E. coli* strains.

1.4.3 Other MGEs

Other chromosomally-borne MGEs include the insertion elements, transposons and integrons. These are segments of DNA that frequently change location in the chromosome. Different categories of such molecules are recognized: (i) insertion sequences with terminal inverted repeats sequences which encode the information for the self-mobilization; (ii) composite transposons flanked on both sides by insertion sequences; (iii) non composite transposons which contain no insertion sequences (Leopold, 2014).
Integrons are well known for their ability to carry and spread antibiotic-resistance cassettes. Recently, super-integrons containing up to eight different cassettes for antibiotic resistance and flanked by insertion sequences targeted by integrases have been described (Cambray *et al.*, 2010, Hall, 2012a, Hall, 2012b).
Plasmids are the only extra-chromosomal MGE. Nonetheless, some of them can also integrate into the host chromosome. Plasmids can exchange genetic information among different strains, thus contributing to the genome evolution by horizontal gene transfer. *E. coli* plasmids can vary in size from few Kb up to more than 150 Kb. Larger plasmids encode virulence factors, such as fimbriae, serine proteases, enterotoxins and haemolysins, among others. Also in plasmids, repeated regions and insertion sequences serve as homologous points for genetic recombination explaining the observed variability in sizes and gene content (Leopold *et al.*, 2014). Several groups of genes have been shown to be originated from plasmid integration into the chromosome and subsequent events of deletion, insertion and stabilization (Brzuszkiewicz *et al.*, 2009).

Horizontal gene transfer operated through MGEs strongly contributes to genome optimization, with potential alterations ranging from the addition of virulence genes to the modification, loss or gain of entire genomic regions. These rapid changes are referred as microevolution and become apparent within a few generations, in the environmental, animal or human reservoir (Leopold *et al.*, 2014).

CHAPTER 2

Aims of the work

The aim of the present work was to investigate the molecular bases of the evolution of pathogenic *E. coli*. The increasing availability of information on the *E. coli* genomes has led to the description of a wealth of mobile genetic elements (MGEs) conveying genes encoding factors involved in the *E. coli* pathogenetic mechanism, but at the same time highlighted that the asset of virulence genes conferring to *E. coli* the full pathogenicity might be by far more complex and is not completely unravelled yet.

The current classification of the *E. coli* pathotypes has been long based on the presence of defined PAIs and virulence genes and the related mechanisms of host colonization or toxin production. However, the boundaries between the different pathogenic types have become less sharp following the observation that part of the accessory DNA, mainly that associated with virulence, seems to overlap different pathotypes causing the concept of *E. coli* pathotypes themselves to be re-considered.

Examples of such an overlapping are the presence of the LEE locus in EPEC and EHEC, or the observed presence of VT-converting phages, hallmark of the VTEC pathotype, into strains belonging to all the other known *E. coli* pathotypes, bringing into question if VTEC are a pathotype in its own right or rather the VT-phage should be considered as being the real pathogen, exploiting the colonization machinery of the pathogenic *E. coli* to generate augmented pathogenicity variants.

Despite the considerable efforts devoted by the scientific community, the evolutionary forces leading to the emergence of new pathogenic clones or variants of *E. coli* are not completely understood. As a matter of fact, *E. coli* genome is very dynamic and evolves continuously through horizontal gene transfer. Indeed, it has been shown that the gain and loss of virulence factors, such as the LEE locus, the VT-converting phages, the Enterohaemolysin-coding gene, among others, have occurred several times and in parallel in separate lineages, thus suggesting that many forces intervened in the *E. coli* convergent evolution (Franz *et al.*, 2014). In this piece of research we focused on VTEC as a model of dynamic pathotype encompassing many established *E. coli* pathogenic groups. We have investigated the distribution of MGEs encoding virulence features in different VTEC subpopulations and studied their evolution through the differentiation of the virulence genes into different allelic variants. Additionally, we studied the different bacteriophages transporting the *vtx* genes into different VTEC types and their role in manipulating the host biology. Finally, we assessed the possibility that the spreading of VT-phages may or may not be subjected to host-related barriers by probing the possible acquisition of such phages by *E. coli* strains with a genetic background different from VTEC.

The work plan has been divided in the following work packages, listed here along with the corresponding detailed objectives:

1. Characterisation of the genetic determinants encoding a novel allelic variant of the Subtilase cytotoxin (SubAB):
    - Analysis of the genetic determinant encoding a new SubAB variant ($SubAB_2$);
    - Study of structure of the MGE carrying the $SubAB_2$-coding genes;
    - Study of the distribution of SubAB-coding genes and their localization onto specific MGEs in VTEC strains from human and animal sources.

      One publication in Chapter 3 (Clin Microbiol Infect. 2013 Mar;19(3):E149-56)

2. Identification of an allelic variant of the virulence-associated gene *toxB* in VTEC serogroups associated with severe human disease:
    - Characterisation of the *toxB* gene in the large virulence plasmid of a VTEC O111 strain;
    - Description of the *toxB* gene allelic variants;
    - Study of the distribution of the two variants in VTEC strains belonging to different serogroups and pathogroups.

      One publication in Chapter 4 (Int J Med Microbiol. 2014 Jul;304(5-6):730-4).

3. Identification and characterisation of VT2-phages present in pathogenic *E. coli* strains causing HUS:

- Identification of a peculiar VT2-phage, Φ-8, present in VTEC O157 strains isolated from severe human disease;
- Functional analysis of the presence of Φ-8 and its influence on the regulation/assembly of the Type III secretion system;
- Comparison between the WGSs of VT2-phage from Enteroaggregative Haemorrhagic *E. coli* (EAHEC) strains and the VT-phages sequences available in the public repositories;
- Identification of peculiar EAHEC VT-phages sequences.
  Two publications in Chapter 5 (Infect Immun. 2014 Jul; 82(7): 3023-32; BMC Genomics. 2014 Jul 8;15:574).

4. Study of the VT-phages host range and the emergence of new pathogenic *E. coli* clones with augmented virulence:

- Determination of the WGSs of VT2-phages from Enteroaggregative Haemorrhagic *E. coli* (EAHEC) and their comparison;
- Analysis of the experimental ability of pathogenic and non-pathogenic *E. coli* strains to be infected with VT2-phages;
- Evaluation of the stable acquisition of VT-phages by *E. coli* strains other than VTEC following experimental infections.
  One publication in Chapter 6 (Front Cell Infect Microbiol. 2014 Jun 20;4:80).

CHAPTER 3

Characterisation of the genetic determinants encoding a novel allelic variant of the Subtilase cytotoxin (SubAB)

Introduction

Subtilase (SubAB) is an AB_5 cytotoxin produced 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 identified in other VTEC serogroups. SubAB is delivered to the host cell endoplasmic reticulum where it causes the inhibition of the protein synthesis, resulting in the unfolded protein response and finally apoptosis. SubAB has been supposed to contribute to the pathogenetic process through a synergistic action with the VTs. In the prototype Subtilase-producing strain, *subAB* operon is located on the large virulence plasmid, named pO113, also carrying the *saa* gene, encoding an autoagglutinating adhesin. Two major allelic variants of the *subAB* genes, termed $subAB_1$ and $subAB_2$, and a minor subtype, termed $subAB_{2\text{-}2}$, have been described so far.

In the present work we determined the complete sequence of the PAI conveying the s$ubAB_2$ genes from a VT-negative *E. coli* strain isolated from a case of diarrhoea in Italy and investigated the distribution of both the major variants in human *E. coli* strains belonging to different pathotypes and in a large collection of LEE-negative VTEC of human and ovine origin.

Results

The entire nucleotidic sequence of the chromosomal locus containing the $subAB_2$ operon 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 a gene encoding an integrase, it harboured the *tia* gene, encoding an invasion determinant first described in an ETEC strain isolated from a human case of disease. Additionally, it presented an ORF showing great sequence similarity with the gene *shiA* of *Shigella.* The study of the distribution of the two *subAB* allelic variants confirmed that *subAB* genes are commonly found in LEE-negative VTEC, showing their presence in 72% and 86% of the strains from human cases of diarrhoea and 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.

Conclusions

The SE-PAI carrying the $subAB_2$ allelic variant has been characterised and fully sequenced. It is inserted in the *E. coli* chromosome in the *pheV*-tRNA locus, which represents a hotspot of integration for genomic islands. The SE-PAI also transports genes, beside *subAB,* which could play a role in the pathogenesis, such as *shiA*, a *Shigella* gene whose product has been described to attenuate the inflammatory response induced in the host upon infection, and *tia*, which encodes an invasion factor of ETEC. The latter genes could play a role in the colonization process by $SubAB_2$-producing *E. coli*. The use of *pheV* as integration site by SE-PAI could provide an explanation for its wide presence in the LEE-negative VTEC strains. In fact it could be the result of a mutual exclusion between this PAI and the LEE locus, which is often integrated in this chromosomal site.

The distribution of the two *subAB* variants has been analysed in a large panel of LEE-negative VTEC strains isolated from human and ovine sources. The results showed that *subAB* was present in the vast majority of LEE-negative VTEC strains isolated from both the ovine source and the human cases of diarrhoea and that $subAB_2$ was the prevalent allele.

Altogether these findings indicate that small ruminants may represent an important reservoir for human infections with VTEC $subAB_2$-producing strains. Moreover, our findings suggest that LEE-negative VTEC strains may express their virulence through the action of additional virulence features including either toxins (SubAB) or factors involved in the colonization mechanism, such as the products of *tia* and *shiA* genes.

Candidate's contributions to the present work:
Laura Grande contributed to the determination of the complete nucleotide sequence of the SE-PAI and helped in the screening of *E. coli* strains held at the Istituto Superiore di Sanità for the presence of *subAB*, *saa* and *tia* genes. Moreover, she revised critically the manuscript for the publication.

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

V. Michelacci[1,2], R. Tozzoli[1], A. Caprioli[1], R. Martínez[3], F. Scheutz[4], L. Grande[1], S. Sánchez[5] and S. 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 and Infection Control, Statens Serum Institut, Copenhagen, Denmark and* 5) *Laboratorio de Enterobacterias, Servicio de Bacteriologia, Centro Nacional de Microbiologia, Instituto de Salud Carlos III, Madrid, Spain*

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 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.

Keywords: STEC, pathogenicity island, subtilase cytotoxin, diarrhoea, small ruminants, PCR
Original Submission: 17 October 2012; **Revised Submission:** 30 November 2012; **Accepted:** 2 December 2012
Editor: F. Allerberger
Article published online: 17 January 2013
Clin Microbiol Infect 2013; **19:** E149–E156
10.1111/1469-0691.12122

Corresponding author: S. Morabito, Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome, Italy
E-mail: stefano.morabito@iss.it

Introduction

Subtilase (SubAB) is an AB_5 toxin produced by certain *Escherichia coli* strains associated with human disease [1]. SubAB is composed of a 35 kDa A subunit displaying a subtilase-like serine protease activity and 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 [4–7].

SubAB has so far been 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 role with Stx [2]. As a matter of fact, SubAB has been shown to induce, in a mouse model, the typical haemolytic uraemic syndrome (HUS)-associated features caused by Stx, such as extensive microvascular damage, and thrombosis and necrosis in the 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 of distinguishing 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 diarrhoeagenic *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 enterotoxin-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, Denmark).

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

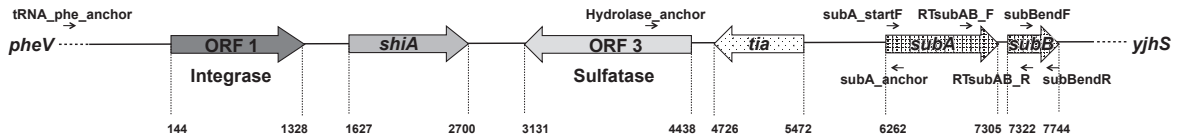


FIG. 1. Map of the Subtilase-Encoding Pathogenicity Island (SE-PAI), spanning the 8058 bp situated between the pheV tRNA locus and the *E. coli* gene yjhS. The ORFs and the annotated genes, together with their positions in the PAI sequence, are shown (GenBank Acc. No. JQ994271). The locations of the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

TABLE 1. PCR primers used in this study for the characterization of the SubAB-encoding PAI, for the detection of *subAB*, *tia* and *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

isolates are part of the culture collection held at the University of Extremadura (Caceres, Spain).

Construction of a genomic library from the 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, USA), following the supplier's instructions. In detail, 200 ng of total DNA were extracted from strain ED 32 (PureGene; Gentra Systems, Big Lake, MN, USA), 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) 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/L NaCl, 2.0 g/L $MgSO_4$ $7H_2O$, 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 1.5×10^9 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. As 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 (Fig. 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 cell assay

To investigate the possible differences in the activity of the two SubAB variants, their ability of inducing CPE onto Vero cell monolayers was 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, was assayed by inoculating serial dilutions, up to 1:1024, onto Vero cell monolayers in microtitre plates. As the ED 186 strain also produced the Stx1, each dilution of the culture supernatant from this strain was neutralized by incubation with an antibody raised against Stx1, at a working titre of 1:200, before inoculation. The VCA was conducted as previously described

TABLE 2. Summary of ORFs in SE-PAI of the 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	sulphatase	sulphatase 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%)

[22] and the CPE observed at 24, 48 and 72 h after inoculation. The *subAB* genes from the two strains were also cloned in the pGEM-T-Easy vector and used to transform the K-12 strain JM109. The *subAB* operons were cloned together with their native promoters to allow the expression of the subtilase. Serial dilutions of overnight culture supernatants of the K-12 strains containing the cloned *subAB* operons were inoculated onto Vero cell monolayers 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 was made through the Glimmer online tool (http://www.ncbi.nlm.nih.gov/genomes/MICROBES/glimmer_3.cgi) and the predicted open reading frames were analysed with the 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 the closest informative matches (Table 2).

All the nucleotidic alignments were carried out 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 Fig. 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 of the $subAB_2$ operon, and the *pheV* tRNA locus, situated 4,725 bp from the end of the *tia* gene (Fig. 1). The latter DNA region contains an unnamed gene encoding a sulphatase, the *shiA* gene, whose product is involved in the attenuation of the *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 the 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 was 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 of 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.

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 previous descriptions for the product of $subAB_2$ allele. Comparable results were also obtained for the two SubAB variants 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 with 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 IS2 element interrupting the *tia* sequence, as has been previously described in strain ED 591 (GenBank Acc. No. FJ664545). Two additional strains (one O91 and one O146)

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)	No. 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 nt[a] (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+[b] (17)	0
EIEC (13)	O121 (1)	0
	O nt[a] (12)	0
UPEC (1)	O102 (1)	0

[a]nt: not tested.
[b]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 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[a]	1[a]
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[b]	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

[a]One single strain was positive for $subAB_1$, $subAB_2$, *saa* and *tia* simultaneously.
[b]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 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[a]	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

[a]The strains have not been tested for the *subAB* variants, *tia* and *saa*, being negative in the *subAB* screening assay (see text for details).

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 IS2 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 the possible animal reservoir of such *E. coli* strains. As 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 (three strains), O76 (one strain) and O176 (one 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 [2]. 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, which we termed SE-PAI, and showed that it is composed of 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 for 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 [1]. 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*$_2$-harbouring 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 the possible sources of such pathogenic *E. coli*. As 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 related to the presence of *tia* in SE-PAI. As a matter of fact, the *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 (i) assess the role of other ruminant species in the epidemiology of these infections and (ii) elucidate whether *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.

Transparency Declaration

No conflict of interest to declare.

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 *et al.* 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 *et al.* 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) serotypes 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 *et al.* 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 *et al.* 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 *et al.* 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 *et al.* 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 *et al.* 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 *et al.* 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.

CHAPTER 4

Identification of an allelic variant of the virulence-associated gene *toxB* in Verocytotoxin-producing *Escherichia coli* (VTEC) serogroups associated with severe human disease

Introduction

All the VTEC virulence factors are encoded on mobile genetic elements. Among them, the gene *toxB* has been first described in the large virulence plasmid of VTEC O157, and later described in similar plasmids carried by VTEC belonging to other serogroups, such as O26 and O145.

toxB is a 9.5 Kb-long gene encoding a protein sharing 28% identity and 47% similarity with the protein encoded by the gene *efa1/lifA* conveyed by the OI-122 island in both VTEC and EPEC.

ToxB has been demonstrated to be involved in the adherence of VTEC O157 to Caco-2 cultured cells promoting the expression of type III secretion system (T3SS)-secreted protein. Some authors suggested that ToxB itself could be translocated trough a T3SS. Accordingly, the distribution of *toxB* in VTEC strains investigated in several studies correlates with the presence of the LEE locus. In the present study we determined for the first time the entire sequence of the *toxB* gene from a VTEC O111 strain isolated from a case of diarrhoea in Italy and observed that this gene is present in two different allelic variants segregating with the VTEC serogroups.

Results

The Whole Genome Sequence (WGS) analysis of a VTEC O111 disclosed the presence, in this strain, of a large virulence plasmid similar to the one carried by a VTEC O145 strain whose sequence was already available in the NCBI nucleotide repository. The sequence data analysis highlighted the presence in the large plasmid of VTEC O111 of a complete 9.5 kb-long intact copy of the *toxB* gene. The comparison of this gene sequence with those present in GenBank showed a 99-100% identity with the *toxB* genes from VTEC strains O26, O145 and O111, while it was only 91% identical at the nucleotide level with the one carried by the VTEC O157 EDL933 strain.

A Real-Time PCR tool able to discriminate between the two *toxB* variants by virtue of their melting temperatures has been designed and used to investigate the relative distribution of the two variants in a panel of 191 VTEC strains belonging to the five SPTs and to multiple serogroups. As expected, *toxB* gene was identified exclusively in VTEC strains belonging to LEE-positive SPTs A and B. The analysis of the curves related to the melting temperature showed that all VTEC O157 strains possessed the gene variant described in the prototype VTEC O157 ($toxB_1$), while all the remaining positive strains belonging to SPT B carried the second variant ($toxB_2$), regardless the serogroup they belonged to.

Conclusions

VTEC are hazardous pathogens causing severe human infections and characterised by the presence of a high number of virulence genes harboured by mobile genetic elements, including the plasmid-borne gene *toxB.*

Some authors suggested a role for ToxB in enhancing the effect of the attaching and effacing histopathological lesion on the gut mucosa of the host. We identified for the first time an intact copy of the *toxB* gene in a VTEC O111 strain and described the existence of two allelic variants of the gene. The screening of a large panel of VTEC for the presence of the *toxB* gene and the identification of its allelic variants, showed that the totality of VTEC O157 strains assayed, constituting the SPT A, possessed the $toxB_1$ variant, while all the positive non-O157 VTEC belonging to the SPT B carried the $toxB_2$ variant. All the VTEC strains belonging to SPT C, D and E and the non-pathogenic *E. coli* tested were negative for the presence of *toxB* gene. This finding substantiates the hypothesis of ToxB contributing to the virulence of the most pathogenic VTEC strains to humans and suggests that the two SPTs A and B may have diverged after the acquisition of the *toxB* gene.

ToxB protein shows a distribution of functional domains similar to the product of the *efa1/lifA* gene, present on the OI-122 in the LEE-positive VTEC, suggesting that the two proteins may have similar functions. Similarly to ToxB, the product of the *efa1/lifA* has been suggested to contribute to the host colonization and such a function has been linked to the inhibition of the host lymphocyte activation. Additionally, the genes

encoding both ToxB and Efa1/lifA share the peculiar size of about 10 Kb, which wouldn't be retained simultaneously in the same cell if not subjected to a strong selective pressure. Although apparently the two proteins could have a redundant function, their concomitant presence in the LEE-positive VTEC belonging to SPTs A and B seem to be related with the prominent virulence of such VTEC groups, while the presence of different alleles of *toxB* gene could be indeed the result of a selection process, with the $toxB_1$ variant present in VTEC O157 strains, which retained the 5' portion of the *efa1/lifA* gene only and could complement the lack, in these strains, of a complete *efa1/lifA* product.

Candidate's contributions to the present work:
Laura Grande contributed to the Real-Time PCR experiments and melting curves data analysis for the study of the distribution of the two *toxB* variants in *E. coli* strains belonging to different seropathotypes. Moreover she helped in the manuscript drafting in the section "Materials and Methods" and revised critically the whole manuscript.

International Journal of Medical Microbiology 304 (2014) 730–734

Contents lists available at ScienceDirect

International Journal of Medical Microbiology

journal homepage: www.elsevier.com/locate/ijmm

Identification of two allelic variants of *toxB* gene and investigation of their distribution among Verocytotoxin-producing *Escherichia coli*

Valeria Michelacci[a,*], Laura Grande[a,b], Rosangela Tozzoli[a], Antonella Maugliani[a], Alfredo Caprioli[a], Stefano Morabito[a]

[a] Istituto Superiore di Sanità, Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare, Rome, Italy
[b] Department of Biology, University 'Roma Tre', Roma, Italy

ARTICLE INFO

Article history:
Received 23 December 2013
Received in revised form 12 May 2014
Accepted 21 May 2014

Keywords:
VTEC
toxB
Virulence
Seropathotype
Plasmid

ABSTRACT

Verocytotoxin-producing *Escherichia coli* (VTEC) are food borne pathogens causing severe human infections. The virulence genes asset of VTEC is complex and has not been completely defined yet. Nonetheless, all the virulence genes described so far have been described as conveyed by mobile genetic elements. A gene, termed *toxB*, has been identified in a large virulence plasmid of VTEC O157, later described in similar plasmids carried by VTEC O26 and O145. In this study we identified for the first time an intact copy of *toxB* gene in a plasmid present in a VTEC O111 strain and observed the existence of two allelic variants of the gene, that we termed $toxB_1$ and $toxB_2$. We investigated the distribution of the two alleles in a panel of VTEC strains belonging to different serogroups and demonstrated that this gene is present only in VTEC serogroups associated with the most severe forms of the infections such as those belonging to the five serogroups O157, O26, O111, O103 and O145 and that the two alleles segregate with the serogroup of the hosting strains. In particular the $toxB_1$ variant was only present in VTEC O157 while the $toxB_2$ allele was present in the remaining four VTEC serogroups.

Introduction

Verocytotoxin (VT)-producing *Escherichia coli* (VTEC) are important zoonotic pathogens (Armstrong et al., 1996). Infection with VTEC results in diarrhea, which in some cases progresses toward more severe forms, such as the hemorrhagic colitis (HC) and the life threatening hemolytic-uremic syndrome (HUS) (Griffin and Tauxe, 1991).

Beside VT production, the pathogenicity of VTEC relies upon the capability of colonizing the intestinal mucosa of the host. The best studied colonization mechanism intervening in VTEC-induced disease is that associated to the "attaching and effacing" (A/E) lesion (Nataro and Kaper, 1998), characterized by the effacement of the microvilli brush border and the intimate adhesion of the bacterium to the enterocyte plasma membrane (Frankel and Phillips, 2008). The formation of A/E lesions is governed by genes conveyed by a pathogenicity island (PAI) termed "locus of enterocyte effacement" (LEE), which encodes a type III secretion system and effector proteins (McDaniel and Kaper, 1997). A scheme has been developed to group VTEC strains on the basis of the serogroup, the presence of the main virulence factors-coding genes, the relative incidence of the infections and the ability to cause severe and epidemic diseases (Karmali et al., 2003). The groups, termed seropathotypes (SPTs), have been labeled with the letters from A to E in a decreasing rank of pathogenicity. SPT A and SPT B comprise the VTEC most frequently involved in outbreaks of severe human infections, HC and HUS. Altogether these SPTs include LEE-positive VTEC belonging to O157, O26, O111, O103, O145 and O121 serogroups with VTEC O157 constituting the SPT A and the remaining grouped in the SPT B (Karmali et al., 2003). SPT C includes VTEC lacking the LEE locus that are occasionally isolated from severe infections but with a very low relative incidence, while SPTs D and E are destined to those VTEC strains belonging to serogroups or possessing a virulence gene asset which have rarely or never been found in VTEC strains isolated from human disease, respectively (Karmali et al., 2003).

In addition to the LEE locus, other virulence factors-coding genes are borne by mobile genetic elements. A large plasmid, termed pO157 (Burland et al., 1998; Makino et al., 1998), is consistently present in VTEC O157 strains (Schmidt et al., 2001) and carries genes such as the *ehxA* gene, which governs the production of the enterohemolysin (Schmidt et al., 1994, 1995) and *toxB* (Makino et al., 1998). The latter is a 9.5 kb-long gene encoding a

* Corresponding author at: Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome, Italy. Tel.: +39 0649902729; fax: +39 0649387101.
E-mail address: valeria.michelacci@iss.it (V. Michelacci).

http://dx.doi.org/10.1016/j.ijmm.2014.05.009
1438-4221/

protein sharing 28% identity and 47% similarity in the amino acidic sequence with the protein encoded by the *efa1/lifA* gene (Tozzoli et al., 2005), conveyed by the OI-122 PAI (Morabito et al., 2003). Similarly to how observed for the product of *efa1/lifA* gene, ToxB has been demonstrated to contribute to the adherence of VTEC O157 to Caco-2 cultured cells by promoting the production and/or the secretion of type III secreted proteins (Stevens et al., 2004; Tatsuno et al., 2000). Further evidences locate the product of *toxB* gene among the effectors translocated through the type III secretion system, suggesting a similar secretion pathway for Efa1/LifA (Deng et al., 2012).

Accordingly, the distribution of *toxB* gene in VTEC investigated in several studies highlighted its strict association only with LEE-positive strains (Bosilevac and Koohmaraie, 2011; Cergole-Novella et al., 2007; Horcajo et al., 2012; Monaghan et al., 2011; Toma et al., 2004; Tozzoli et al., 2005).

The presence of *toxB* gene has been recently identified in the published sequence of the large virulence plasmids of two O26 (Fratamico et al., 2011; Venturini et al., 2010) and one O145:NM VTEC strains (Yan et al., 2012).

In the present study we determined the entire sequence of the *toxB* gene from an O111 VTEC strain isolated from a case of diarrhea in Italy in 2011. Moreover, we observed the existence of two different allelic variants of *toxB* gene, one specific for VTEC O157 strains (SPT A) and the other associated with the VTEC strains belonging to the SPT B.

Materials and methods

Bacterial strains

The VTEC strains analyzed in this study are part of the culture collections held at the Istituto Superiore di Sanità (Rome) and have been isolated in Italy from human cases of infections and animal sources in the period spanning years 1988 and 2008. A total of 191 VTEC epidemiologically unrelated strains have been assayed, comprising 51 O157 strains, 35 O26, six O103 strains, 32 O111, four O121, eight O145, ten O113 and five O91. Forty strains of animal origin belonging to different serogroups (Table 1) were also included in the study. The VTEC O157 reference strains EDL 933 (Wells et al., 1983) and RIMD0509952/VT2 Sakai (Watanabe et al., 1996) have been included as controls together with 27 non-pathogenic *E. coli* human isolates part of the reference ECOR collection (Ochman and Selander, 1984).

Sequencing of the O111 VTEC strain ED 734 and bioinformatics analysis

The whole genome sequencing of the O111 VTEC strain ED 734 has been outsourced to IGA Technology Services S.R.L. (Udine, Italia) and performed with the HiSeq2000 sequencer by Illumina Inc. (San Diego, CA, USA) using a paired-end sequencing protocol (2× 100 bp). The FASTQ files have been assembled in contigs through the de novo assembly tool part of the CLC Genomics Workbench program (CLC Bio, Aarhus, Denmark). The contigs obtained have been automatically annotated by using PROKKA program (http://www.vicbioinformatics.com/software.prokka.shtml) and compared on MAUVE software (Darling et al., 2010) with the complete sequences from plasmids pO157-EDL933, pO26-Vir and pO145 already available on GenBank database (Acc. Nos. AF074613, NC_012487, and HM138194, respectively). The coding sequence of *toxB* gene from ED 734 strain has been deposited in GenBank database with the Acc. No. JQ906697.

Multiple alignments of *toxB* gene sequences have been performed through ClustalW2 webserver located at EMBL-EBI website (http://www.ebi.ac.uk/Tools/msa/clustalw2/) (Larkin et al., 2007), while MEGA 5 bioinformatics software (Tamura et al., 2011) has been used to graph the results in a phylogenetic tree.

Real time-PCR experiments and melting curve analysis

The primers toxB_F (5′-tggtaatcatggggcaaaat-3′) and toxB_R (5′-tcaccgctaggcaaatatatca-3′) used in the real time-PCR reactions have been designed on conserved regions and able to amplify part of the *toxB* genes borne on the pO157 plasmid of the O157:H7 EDL 933 strain (Acc. No. AF074613), the pO26-Vir plasmid of the O26:H11 H30 strain (Acc. No. NC_012487), the pO145 of the O145:NM 83-75 strain (Acc. No. HM138194) and the large virulence plasmid of the O111 ED 734 strain, identified in this study (Acc. No. JQ906697).

The predictive analysis of the melting temperatures for the amplified products has been carried out in silico by using the melting temperature analysis tool available at the Integrated DNA Technologies webserver (http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/). The real time-PCR reactions have been performed in 15 μl by the use of the SSO-Fast EvaGreen master mix (BioRad, Hercules, CA, USA), 2 μl of a lysed culture as template and the toxB_F and toxB_R primers at a concentration of 250 nM each. The reactions have been processed on a Rotorgene 6000 instrument (Corbett Life science, Concorde, Australia) through 40 cycles of amplification (10 s at 95 °C; 45 s at 60 °C) followed by a slow melting step operated by gradually increasing the temperature starting from 50 °C up to 95 °C, at a rate of 0.5 °C/s, and continuously monitoring the decreasing fluorescence. The resulting melting curves can be mathematically derived to eventually result in peaks corresponding to the denaturation temperatures of the amplified products.

Results

Identification of toxB gene in the plasmid of the O111 VTEC strain ED 734

The raw data deriving from the whole genome sequencing of the O111 ED 734 VTEC strains were de novo assembled for a total of 5,222,087 bp in 132 contigs (N50 score 210345), which were automatically annotated and aligned to the complete sequences of the plasmids derived from VTEC and available in GenBank, with the aim of identifying and characterizing the plasmids possibly present in the strain ED 734.

This analysis disclosed the presence of a virulence plasmid, which was termed pO111-ED 734. Such a plasmid was very similar to the pO145 harbored by the 83-75 VTEC O145 strain already sequenced (Acc. No. HM138194) (Fig. 1), even though some pO145 DNA regions appeared to be absent in pO111-ED 734 (Fig. 1). Such regions were made up by insertion sequences and transposable elements and could have been lost following excision events. Nevertheless, the observed absence of these mobile genetic elements from the pO111-ED 734 could also result from a failure in the matching process during the assembly, therefore further work is in progress to close up such gaps. In any case, the analysis highlighted the presence of a complete copy of *toxB* gene (9.5 kb) in a 19 kb-long contig, perfectly matching the corresponding region comprising *toxB* in the pO145 sequence (Fig. 1).

The nucleotidic sequence of *toxB* gene from the ED 734 strain was deposited in GenBank database with the Acc. No. JQ906697 and was further analyzed.

Identification of the two allelic variants of toxB gene

We performed in silico comparison of *toxB* nucleotidic sequences present in the GenBank database with that determined

Table 1
Screening of the *E. coli* strains for the presence of *toxB* gene and distribution of the allelic variants of the gene.

Seropathotype	Serogroup	No. of strains	No. *toxB*-positive strains	*toxB* variant
A	O157	51	51	$toxB_1$
B	O26	35	27	$toxB_2$
	O103	6	0	–
	O111	32	2	$toxB_2$
	O121	4	4	$toxB_2$
	O145	8	4	$toxB_2$
C	O91	5	0	–
	O113	10	0	–
D and E	O1	1	0	–
	O2	2	0	–
	O8	4	0	–
	O11	1	0	–
	O15	1	0	–
	O18	2	0	–
	O25	1	0	–
	O39	1	0	–
	O73	1	0	–
	O82	1	0	–
	O75	2	0	–
	O88	4	0	–
	O101	3	0	–
	O109	1	0	–
	O110	1	0	–
	O126	1	0	–
	O139	2	0	–
	O152	1	0	–
	O171	3	0	–
	O rough	2	0	–
	O NT[a]	5	0	–
ECO R collection	Various	27	0	–
Total		218	88	

[a] Not typable.

in the plasmid pO111-ED 734. A ClustalW2 analysis was conducted using the *toxB* sequences from the VTEC strain O157 EDL 933 (Acc. No. AF074613), the O26 strain H30 (Acc. No. NC_012487), the O145 strain 83-75 (Acc. No. HM138194) and the O111 strain ED 734 (Acc. No. JQ906697). Such an analysis revealed that the *toxB* genes harbored by the large virulence plasmids of the VTEC O26, O145 and O111 strains are nearly identical (99–100% of sequence identity), while the copy harbored by the O157 strain EDL 933 shared only 91% identity at the nucleotidic sequence level with the remaining three genes analyzed (Fig. 2).

Similar results were obtained by performing the same analysis using the deduced amino acidic sequences of ToxB protein. The latter alignments returned 100% and 99% (3165/3166) of sequence identity between the ToxB protein from the pO111-ED 734 and those from the O145 strain 83-75 and the O26 strain H30, respectively, and 90% (2851/3168) of sequence identity and 95% (3010/3168) of positive substitutions with ToxB from the O157 strain EDL 933.

These results indicate the existence of at least two distinct allelic variants of the *toxB* gene, with one being associated with VTEC O157

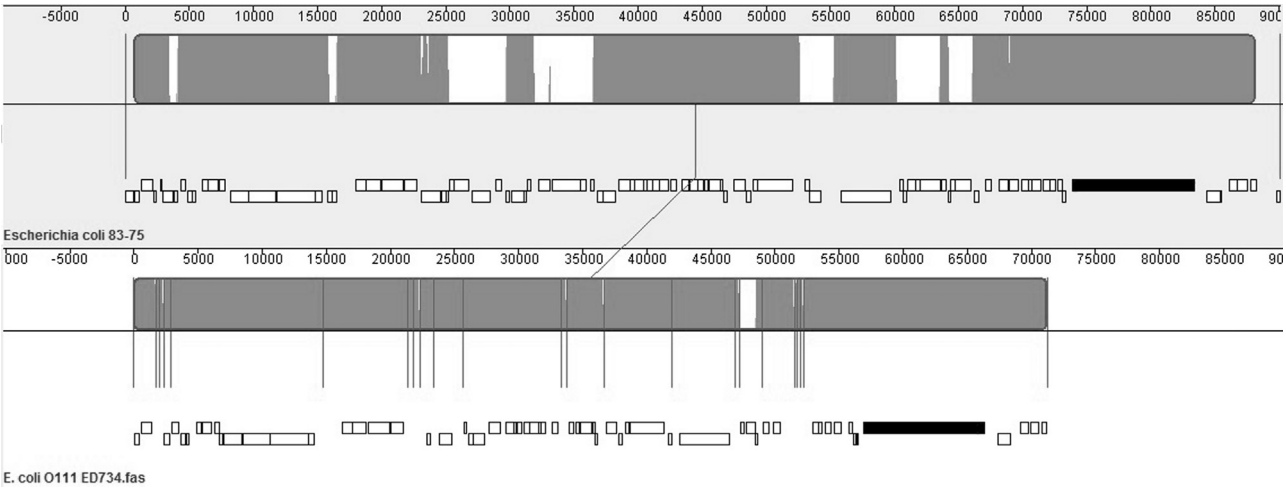


Fig. 1. Progressive alignment of pO145 (Acc. No. HM138194) (on the upper level) and the pO111-ED 734 plasmid harbored by the O111 ED 734 sequenced strain (lower level) using the MAUVE software. The sequences are outlined as blocks with the conserved regions colored in gray, while the white regions represent strain-specific stretches. The position of *toxB* gene is marked in black in the brick-shaped diagram underlying the two blocks and indicating the open reading frames.

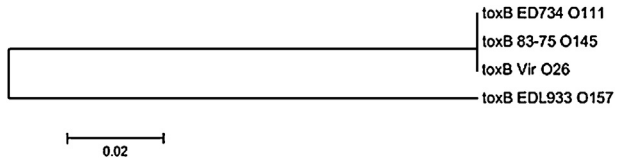


Fig. 2. ClustalW2 alignment of the available complete nucleotidic sequences of *toxB* genes from VTEC strains belonging to different serogroups. The scale represents a number of substitutions per base per indicated horizontal distance.

strains, which we denominated $toxB_1$; and a second one, that we termed $toxB_2$, present in VTEC strains belonging to the serogroups O26, O145 and O111.

Detection of the two allelic variants of toxB by real time-PCR and melting analysis

We deployed a tool aiming at identifying the two allelic forms of *toxB* gene and used it to investigate their relative distribution in a panel of VTEC strains belonging to multiple serogroups. The toxB_F/toxB_R primer pair, allowing the amplification of both the variants, was used for the amplification of a 251 bp product, characterized by a GC content of 29.5% and 33.9% for the $toxB_1$ and the $toxB_2$ alleles, respectively. This difference, in turn, allows the discrimination between the two amplification products by virtue of the different temperature peaks associated to the melting curves related to the two amplicons, resulting in melting temperatures peaks of 77 °C for $toxB_1$ and 79 °C for $toxB_2$ (Fig. 3).

Distribution of toxB gene and relative presence of its alleles in VTEC strains belonging to the different seropathotypes

A total of 191 VTEC strains belonging to the different seropathotypes were tested for the presence of the two *toxB* alleles. As expected, only the strains belonging to the most pathogenic seropathotypes (A and B) proved positive for the presence of *toxB*. The presence of the *toxB* gene was detected in 100% (51/51) of the O157 strains assayed (SPT A). As far as the SPT B strains were concerned, the presence of the gene was observed in the four O121 strains assayed, in 77% of the O26 strains (27/35), and in half of the O145 strains tested (4/8), while only 2 out of the 32 VTEC O111 strains and none of the six O103 strains assayed were positive in the *toxB*-specific PCR (Table 1). Finally, none of the 55 strains belonging to seropathotypes C, D and E tested were positive for the presence of *toxB*, as well as none of the 27 non-pathogenic *E. coli* strains belonging to the ECOR collection. The analysis of the curves related to the melting temperature of the *toxB* amplicons showed that all the 51 O157 strains harbored the $toxB_1$ allele with the melting peak pointed at 77.00 ± 0.15 °C, while all the remaining

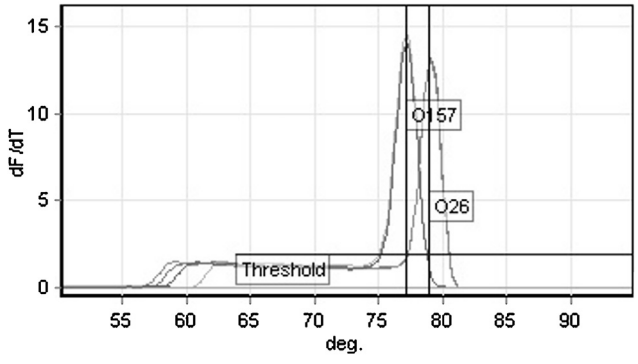


Fig. 3. Analysis showing the peaks corresponding to the melting temperature of the amplification products obtained from $toxB_1$ (77 °C) and $toxB_2$ (79 °C).

37 positive strains belonging to SPT B possessed the $toxB_2$ variant, whose melting peak was at 79.00 ± 0.15 °C (Table 1).

Discussion

VTEC are hazardous pathogens causing severe human infection characterized by the presence of a high number of virulence genes harbored by mobile genetic elements, including the plasmid-borne gene *toxB* (Fratamico et al., 2011; Tozzoli et al., 2005; Yan et al., 2012).

In this study we identified for the first time an intact copy of *toxB* gene in the large virulence plasmid present in a VTEC O111 and investigated the distribution of this gene in the different VTEC seropathotypes by assaying a panel of 191 epidemiologically unrelated VTEC strains. The presence of *toxB* was observed only in strains belonging to SPTs A and B. This observation is in line with the presence of this gene in the sequence of the plasmids of VTEC O26, VTEC O157 and VTEC O145 available in GenBank (Acc. Nos. NC_012487, AF074613, and HM138194, respectively) and with its absence in the large virulence plasmid of strains belonging to SPT C such as the VTEC O113 strain EH41 (Acc. No. NC_007365) (Srimanote et al., 2002). This finding substantiates the hypothesis for a role of ToxB in the pathogenesis of VTEC infections caused by SPT A and some SPT B VTEC strains. Interestingly, the SPT A and B VTEC all share the presence of the LEE Pathogenicity Island, supporting a role of ToxB in enhancing the effect of the attaching and effacing hystopathological lesion on the gut mucosa of the host, as some authors have suggested (Stevens et al., 2004; Tatsuno et al., 2000).

We found that *toxB* gene was absent in some of the SPT B strains analyzed. This finding might be put into relation with the mobile nature of the virulence genes characterized in the VTEC stains. As a matter of fact, the *toxB* gene can be considered as a mobile genetic element in its own right, being surrounded at both ends by IS3-like elements (Tozzoli et al., 2005). Therefore, *toxB* might be capable of self-movement and exciding itself from the large virulence plasmid, which was still present in 40 out of the 48 SPT B *toxB*-negative strains as assessed by detecting the presence of the *ehxA* gene (Schmidt et al., 1995), which is considered as a hallmark for the presence of the large virulence plasmid (data not shown). The remaining eight strains were also negative for the *ehxA* gene, indicating that in these isolates the whole plasmid was lost.

By comparing the published sequences of the virulence plasmids of VTEC O157, O26 and O145 with that of the pO111 determined in this study, we observed that *toxB* gene exists into at least two allelic variants, that we termed $toxB_1$ and $toxB_2$. The two alleles segregated with the serogroups of the VTEC strains and in particular, $toxB_1$ form was present in the pO157, while the $toxB_2$ was harbored by the plasmids sequenced from the non-O157 VTEC (Acc. Nos. NC_012487, AF074613, and HM138194).

The distribution of the two *toxB* alleles has been studied by analyzing the melting curves of the real time-PCR amplicons obtained from the screening of the panel of VTEC strains and confirmed the distribution of the two alleles, $toxB_1$ and $toxB_2$, with seropathotypes A and B, respectively (Table 1).

We speculated on the distribution of the two *toxB* allelic forms in the light of the function of the ToxB protein. A similar distribution of functional domains has been observed between ToxB and the product of the *efa1/lifA* gene (Tozzoli et al., 2005). This observation suggests the two proteins may have similar functions. At the same time, the conservation of such two long genes encoding similar functions is puzzling from an evolutionary view.

The *efa1/lifA* gene is harbored on the PAI OI-122 and encodes a factor involved in the host immunomodulation (Klapproth et al., 2000). The Efa1/lifA protein has been also described as a colonization factor contributing to the adherence of EPEC to cultured cell

lines (Nicholls et al., 2000). Additionally, its influence on the colonization of cattle has been demonstrated (Deacon et al., 2010; Stevens et al., 2002).

Similar functions have been attributed to the *toxB* product (Stevens et al., 2004; Tatsuno et al., 2000) suggesting that the two proteins could have a redundant effect.

Interestingly, the *efa1/lifA* gene is present in its entire form only in strains belonging to SPT B (Morabito et al., 2003), while it is truncated in O157 strains, which only harbor about 800 bp at the 5′ of the coding sequence (Morabito et al., 2003). The polymorphisms individuating the two *toxB* alleles might therefore reflect differences in the functionality of the gene products, with $toxB_1$ encoding a factor able to complement the lack of a complete Efa1/LifA product.

Further work is needed to clarify this issue and to disclose the exact role of ToxB in the pathogenesis of the disease induced by VTEC.

Acknowledgements

This work was partially supported by a grant from Bio-Rad Laboratories (Hercules, USA).

References

Armstrong, G.L., Hollingsworth, J., Morris Jr., J.G., 1996. Emerging foodborne pathogens: *Escherichia coli* O157:H7 as a model of entry of a new pathogen into the food supply of the developed world. Epidemiol. Rev. 18, 29–51.

Bosilevac, J.M., Koohmaraie, M., 2011. Prevalence and characterization of non-O157 Shiga toxin-producing *Escherichia coli* isolates from commercial ground beef in the United States. Appl. Environ. Microbiol. 77, 2103–2112.

Burland, V., Shao, Y., Perna, N.T., Plunkett, G., Sofia, H.J., Blattner, F.R., 1998. The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. Nucleic Acids Res. 26, 4196–4204.

Cergole-Novella, M.C., Nishimura, L.S., Dos Santos, L.F., Irino, K., Vaz, T.M., Bergamini, A.M., Guth, B.E., 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.

Darling, A.E., Mau, B., Perna, N.T., 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS ONE 5, e11147.

Deacon, V., Dziva, F., van Diemen, P.M., Frankel, G., Stevens, M.P., 2010. Efa-1/LifA mediates intestinal colonization of calves by enterohaemorrhagic *Escherichia coli* O26: H- in a manner independent of glycosyltransferase and cysteine protease motifs or effects on type III secretion. Microbiology 156, 2527–2536.

Deng, W., Yu, H.B., de Hoog, C.L., Stoynov, N., Li, Y., Foster, L.J., Finlay, B.B., 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 – MCP 11, 692–709.

Frankel, G., Phillips, A.D., 2008. Attaching effacing *Escherichia coli* and paradigms of Tir-triggered actin polymerization: getting off the pedestal. Cell. Microbiol. 10, 549–556.

Fratamico, P.M., Yan, X., Caprioli, A., Esposito, G., Needleman, D.S., Pepe, T., Tozzoli, R., Cortesi, M.L., Morabito, S., 2011. The complete DNA sequence and analysis of the virulence plasmid and of five additional plasmids carried by Shiga toxin-producing *Escherichia coli* O26:H11 strain H30. Int. J. Med. Microbiol. – IJMM 301, 192–203.

Griffin, P.M., Tauxe, R.V., 1991. The epidemiology of infections caused by *Escherichia coli* O157:H7, other enterohemorrhagic *E. coli*, and the associated hemolytic uremic syndrome. Epidemiol. Rev. 13, 60–98.

Horcajo, P., Dominguez-Bernal, G., de la Fuente, R., Ruiz-Santa-Quiteria, J.A., Blanco, J.E., Blanco, M., Mora, A., Dahbi, G., Lopez, C., Puentes, B., Alonso, M.P., Blanco, J., Orden, J.A., 2012. Comparison of ruminant and human attaching and effacing *Escherichia coli* (AEEC) strains. Vet. Microbiol. 155, 341–348.

Karmali, M.A., Mascarenhas, M., Shen, S., Ziebell, K., Johnson, S., Reid-Smith, R., Isaac-Renton, J., Clark, C., Rahn, K., Kaper, J.B., 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.

Klapproth, J.M., Scaletsky, I.C., McNamara, B.P., Lai, L.C., Malstrom, C., James, S.P., Donnenberg, M.S., 2000. A large toxin from pathogenic *Escherichia coli* strains that inhibits lymphocyte activation. Infect. Immun. 68, 2148–2155.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.

Makino, K., Ishii, K., Yasunaga, T., Hattori, M., Yokoyama, K., Yutsudo, C.H., Kubota, Y., Yamaichi, Y., Iida, T., Yamamoto, K., Honda, T., Han, C.G., Ohtsubo, E., Kasamatsu, M., Hayashi, T., Kuhara, S., Shinagawa, H., 1998. Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic *Escherichia coli* O157:H7 derived from Sakai outbreak. DNA Res.: Int. J. Rapid Publ. Rep. Genes Genomes 5, 1–9.

McDaniel, T.K., Kaper, J.B., 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.

Monaghan, A., Byrne, B., Fanning, S., Sweeney, T., McDowell, D., Bolton, D.J., 2011. Serotypes and virulence profiles of non-O157 Shiga toxin-producing *Escherichia coli* isolates from bovine farms. Appl. Environ. Microbiol. 77, 8662–8668.

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, 3343–3348.

Nataro, J.P., Kaper, J.B., 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. 11, 142–201.

Nicholls, L., Grant, T.H., Robins-Browne, R.M., 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.

Ochman, H., Selander, R.K., 1984. Standard reference strains of *Escherichia coli* from natural populations. J. Bacteriol. 157, 690–693.

Schmidt, H., Beutin, L., Karch, H., 1995. Molecular analysis of the plasmid-encoded hemolysin of *Escherichia coli* O157:H7 strain EDL 933. Infect. Immun. 63, 1055–1061.

Schmidt, H., Karch, H., Beutin, L., 1994. The large-sized plasmids of enterohemorrhagic *Escherichia coli* O157 strains encode hemolysins which are presumably members of the *E. coli* alpha-hemolysin family. FEMS Microbiol. Lett. 117, 189–196.

Schmidt, H., Bitzan, M., Karch, H., 2001. Pathogenic aspects of Shiga toxin-producing *E. coli* infections in humans. In: Duffy, G., Garvey, P., McDowell, D.M. (Eds.), Verocytotoxigenic *Escherichia coli*. Food & Nutrition Press Inc., Trumbull, CT.

Srimanote, P., Paton, A.W., Paton, J.C., 2002. Characterization of a novel type IV pilus locus encoded on the large plasmid of locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. Infect. Immun. 70, 3094–3100.

Stevens, M.P., Roe, A.J., Vlisidou, I., van Diemen, P.M., La Ragione, R.M., Best, A., Woodward, M.J., Gally, D.L., Wallis, T.S., 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.

Stevens, M.P., van Diemen, P.M., Frankel, G., Phillips, A.D., Wallis, T.S., 2002. Efa1 influences colonization of the bovine intestine by Shiga toxin-producing *Escherichia coli* serotypes O5 and O111. Infect. Immun. 70, 5158–5166.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.

Tatsuno, I., Kimura, H., Okutani, A., Kanamaru, K., Abe, H., Nagai, S., Makino, K., Shinagawa, H., Yoshida, M., Sato, K., Nakamoto, J., Tobe, T., Sasakawa, C., 2000. Isolation and characterization of mini-Tn5Km2 insertion mutants of enterohemorrhagic *Escherichia coli* O157:H7 deficient in adherence to Caco-2 cells. Infect. Immun. 68, 5943–5952.

Toma, C., Martinez Espinosa, E., Song, T., Miliwebsky, E., Chinen, I., Iyoda, S., Iwanaga, M., Rivas, M., 2004. Distribution of putative adhesins in different seropathotypes of Shiga toxin-producing *Escherichia coli*. J. Clin. Microbiol. 42, 4937–4946.

Tozzoli, R., Caprioli, A., Morabito, S., 2005. Detection of toxB, a plasmid virulence gene of *Escherichia coli* O157, in enterohemorrhagic and enteropathogenic *E. coli*. J. Clin. Microbiol. 43, 4052–4056.

Venturini, C., Beatson, S.A., Djordjevic, S.P., Walker, M.J., 2010. Multiple antibiotic resistance gene recruitment onto the enterohemorrhagic *Escherichia coli* virulence plasmid. FASEB J.: Off. Publ. Feder. Am. Soc. Exp. Biol. 24, 1160–1166.

Watanabe, H., Wada, A., Inagaki, Y., Itoh, K., Tamura, K., 1996. Outbreaks of enterohaemorrhagic *Escherichia coli* O157:H7 infection by two different genotype strains in Japan, 1996. Lancet 348, 831–832.

Wells, J.G., Davis, B.R., Wachsmuth, I.K., Riley, L.W., Remis, R.S., Sokolow, R., Morris, G.K., 1983. Laboratory investigation of hemorrhagic colitis outbreaks associated with a rare *Escherichia coli* serotype. J. Clin. Microbiol. 18, 512–520.

Yan, X., Fratamico, P.M., Needleman, D.S., Bayles, D.O., 2012. DNA sequence and analysis of a 90.1-kb plasmid in Shiga toxin-producing *Escherichia coli* (STEC) O145:NM 83-75. Plasmid 68, 25–32.

CHAPTER 5

Identification and characterisation of VT2-phages present in Verocytotoxin-producing *Escherichia coli* strains causing the Haemolytic-Uremic Syndrome

Introduction

The ability to produce the Verocytotoxins (VTs) is the striking virulence feature of a class of pathogenic *E. coli* referred to as the Verocytotoxin-producing *Escherichia coli* (VTEC). VT-coding genes are conveyed by bacteriophages and VTEC have been reported as being isolated from cases of human disease displaying a wide variety of symptoms, ranging from mild diarrhoea to the life-threatening haemolytic uremic syndrome (HUS). We focused our attention on VT-phages present in VTEC strains causing HUS, the most severe form of disease associated to infection with these bacterial pathogens. In particular we considered VTEC O157, the prototype of Enterohaemorrhagic *E. coli*, and the VTEC strains that have recently broke the scene following the large outbreak of infections in Germany in 2011 and termed Enteroaggregative Haemorrhagic *E. coli* (EAHEC).

We have identified and characterised a peculiar VT-phage carried by VTEC O157 strains, which is significantly associated to strains causing human illness but less represented in VTEC O157 belonging to the Phage Types (PT) principally circulating in the animal reservoir and described that its presence has an effect on the regulation of genes present in the LEE locus. Additionally, we have determined the complete genome sequence of the VT-phage carried by the first EAHEC strain ever described and compared it with the whole genome sequences (WGSs) of the other VT-phages from the EAHEC strains available in the public repositories and described that the phages conveying the *vtx2*-genes in EHAEC strains belonged to at least two different types.

Publication n. 1 (Infect Immun. 2014 Jul; 82(7): 3023-32).

Results

We have chosen to start our investigation by selecting VTEC O157 strains belonging to the PTs described as characterising strains circulating in the animal reservoir and isolated from human infections. Ten VTEC O157 strains belonging to PT21 isolated from animals and ten humans strains belonging to PT8 have been compared at the whole genome level by DNA-microarray hybridization using slides containing the whole complement of ORFs from VTEC O157 strains and *E. coli* K-12 strains. The comparative genomics hybridization analysis showed that the major differences between the two groups of VTEC strains were in two DNA sequences of the VT2-converting bacteriophage, with the PT21 strains showing a VT2-phage similar to the one carried by the VTEC O157 reference genome, EDL933. One of the two polymorphic regions identified is located between the *gam* and *cII* genes and contains genes whose products are responsible for the switch between the lytic and lysogenic cycles. The other region was that between the genes *roi* and *s*, activated upon the induction of the lytic cycle. In accordance with the microarray hybridisation results, the long PCR amplification of the *gam/cII* region of phage present in the PT21 strains gave an amplicon of 6.1 Kb, the size expected for the VT2-phage BP933W present in the reference strain EDL933, while the same PCR yielded a 4.9 Kb size amplicon in all the PT8 strains. The VT2-phage possessing this region in PT8 strains was termed Φ-8.

The *gam/cII* region of Φ-8 is 99% homologue to the same one present in the VT2-phage of the VTEC O157 strain isolated during an outbreak that occurred in Japan in 1996 and was detected, using a specifically deployed PCR tool, in a proportion of VTEC O157 isolated from human disease higher than that observed in the animal isolates ($P<0.0001$), regardless the PT. The same region was also identified in all the assayed VTEC O157 of PT21/28 (30/30), the PT most frequently isolated from human infections in the United Kingdom and described to be frequent in animals presenting high-level excretion of VTEC O157 with their faeces.

To ascertain if differences exist in the ability to produce VTs of the two types of phage, the *vtx2A* mRNA has been quantified in cultures of VTEC O157 strains carrying Φ-8 or the BP933W-like phage. The experiments did not show significant differences between the two groups, suggesting that the reason for the

observed association of Φ-8 with VTEC O157 from human disease must reside elsewhere. We have also explored the possibility that the presence of Φ-8 could influence the regulation of other virulence determinants, as it was hypothesised in other systems. In particular, we have evaluated if the presence of this phage could have an effect on the production and assembly of the Type III Secretion System (T3SS), a key virulence feature of VTEC O157 strains involved in the attaching and effacing lesion, by analysing its influence on the transcription of the genes encoding the T3SS components. This was evaluated using a construct made up of the LEE1 operon master promoter, which regulates the transcription of the T3SS genes, controlling the expression of the Green Fluorescent Protein gene as a reporter. Such a construct was co-transformed into an *E. coli* K-12 strain together with another plasmid carrying the *gam/cII* region of Φ-8 or the same region from the control *vtx2*-phage BP933W. This approach allowed us observing a marked down-regulation of the LEE1 promoter activity caused by the presence of the Φ-8 specific region in comparison with how observed when, in the same system, the *gam/cII* region of the BP933W phage was cloned.

Accordingly, we could observe that the wild type VTEC O157 strains carrying Φ-8 or the BP933W showed differences in the amount of the secreted EspD protein, part of and secreted through the T3SS, as evaluated by Western Blot analysis of the supernatant of cultures of the two VTEC O157 groups, with the strains carrying Φ-8 showing lower levels of secreted EspD.

Conclusions

The hypothesis that different VTEC O157 clones could be characterised by higher virulence or more efficiently transmitted to the human host has been formulated by several authors, based on different observations including a non-random distribution of PTs between strains isolated from the animal reservoir, such as PT21, and human disease, which mainly include PT2, PT8 and PT 21/28. We compared the whole set of coding regions from VTEC O157 strains belonging to PT8 and PT21 and identified a region of the VT2-phage, comprised between *gam* and *cII* genes, accounting for the most significant difference between the two groups. Such a region is normally involved in the switch between the lysogenic state and the lytic cycle of the lambdoid phages. Interestingly, VT2-phages with a *gam/cII* region similar to that of the VT2-phages of the PT8 strains, that we termed Φ-8, were significantly associated with VTEC O157 strains isolated from human diseases in Italy, regardless their PT. Moreover, the same sequence could be detected in the VT2-phages of VTEC O157 strains belonging to PT21/28 and isolated from super-shedders cattle in the United Kingdom.

The reason for the observed strong association between Φ-8 phage and VTEC O157 from human disease appears not to be due to an increased ability of the former phage to produce higher amounts of VT2, as VTEC O157 strains carrying different types of phages showed comparable levels of toxin's mRNA, at least in the conditions used in the laboratory.

Besides the production of VTs, the ability to colonize the intestinal mucosa with the T3SS-mediated attaching and effacing (A/E) lesion is considered to be pivotal to the pathophysiology of VTEC O157-induced disease. We have explored the possibility that the presence of Φ-8 could have an influence on the ability of VTEC O157 to cause the A/E lesion to the enterocyte and discovered, by using an *E. coli* K-12 in vitro system, that it seems to down-regulate the main promoter of the LEE locus, in turn tuning the expression of the genes involved in the production and assembly of the T3SS. Accordingly, Φ-8 positive wild type VTEC O157 strains also showed lower levels of the secreted EspD, a T3SS marker protein.

A similar result on the levels of the secreted EspD and the concomitant presence of Φ-8 has been also demonstrated in VTEC O157 strains isolated from super-shedder animals. These results strengthen the hypothesis of a role for Φ-8 in influencing the colonization of the host and may have implications for the persistence of VTEC O157 hosting this VT-phage in the animal reservoir. Super-shedding is important for VTEC O157 to establish and maintain in the environment and the observed regulation of T3SS exerted by

Φ-8 may favour the high excreting phenotype in cattle, boosting the presence of VTEC O157 in the farms and eventually causing an increased exposure of humans to VTEC O157, explaining the observed over-representation of this phage in VTEC O157 strains isolated from human infections.

Publication n. 2 (BMC Genomics. 2014 Jul 8;15:574).

Results

The genome of the VT-phage of the first EAHEC strain, isolated in the '90s during a small HUS outbreak, was fully determined and annotated. The genome of the phage, termed Phi-191, was about 61 Kb long and included 87 predicted coding sequences and three transfer RNAs (tRNAs). Phi-191 genome included the genes *cI* and *cro*, two of the main regulators of the lambdoid phages, while the other genes typically composing the regulatory repertoire of these phages such as *cII, cIII, N*, *Ea10* and *gam* seemed to be missing. A BLAST search was run against all the phage sequences present in the GenBank repository at NCBI and returned a 99% identity with that of the VT-phage from the EAHEC strain responsible of the German outbreak in 2011 and a 91% with that from the VTEC O103:H25 strain that caused a severe HUS outbreak in Norway in 2006. Furthermore, Phi-191 sequence was highly related to the VT-phage sequence from two EAHEC O104:H4 strains isolated from HC cases in Georgia in 2009. The other scores included hits displaying variable query coverage with a number of VT-phages from several VTEC strains. Additional comparisons made against the whole genome sequence of an EAHEC O111:H21 isolated from an HUS case in northern Ireland in 2012 present in the short reads archive at the GenBank showed that Phi-191 phage seems to be largely divergent from the VT2-phage present in this strain.

The alignment of the Phi-191 sequence with those from other EAHEC or typical VTEC highlighted the presence of two sequence blocks that seemed to be peculiar to the EAHEC VT2-phages and divergent from the same regions of the other VTEC associated phages. The first fragment was 1,500 bp long and included part of a gene encoding a lysozyme, the lysis protein S and two hypothetical proteins. The second fragment was 900 bp long and belonged to a gene apparently coding for a phage tail fiber. The sole 900 bp long fragment could also be detected in the genome of the O111:H21 EAHEC isolated in Northern Ireland in 2012.

Conclusions

EAHEC strains belonging to at least six different serotypes have been described in the last twenty years. It has been proposed that EAHEC strains arose from typical Enteroaggregative *E. coli* strains that stably acquired a VT2-phage. We determined the WGS of the VT2-phage isolated from the first EAHEC strain ever described and compared it with those of other EAHEC and VTEC VT-phages, with the aim of understanding the events underlying the emergence of such a new pathotype. We could observe that the VT2-phage from the first EAHEC strain isolated in the '90s and from the strain that have caused the German outbreak of 2011 are identical although isolated in a 20-years' time span period and despite the described huge variability of VT-phages. This finding suggests that this phage might be kept under a strong selective pressure, raising question on its provenience. Recent papers propose that the origin of the VT2-phages of EAHEC could be in VTEC strains originating from cattle. This hypothesis seems to provide a meaningful picture for the emergence of EAHEC, but it is in contrast with the mentioned stability of the EAHEC VT-phage genome. As a matter of fact, VT-phages present in VTEC inhabiting the animal reservoir are abundant in number and types facilitating the frequent recombination and exchange of genetic traits and explaining the described high diversity of these microorganisms. On the other hand, the observation that the genome of the VT2-phage present in most of the six the EAHEC serotypes described so far remained untouched for more than 20 years, seems to indicate that these phages are maintained into a reservoir that applies a strong selection on them or where the VT-phages might not be abundant, making the genetic drift as the unique engine for their evolution.

Additionally, the VT2-phage from the O111:H21 EHAEC strain isolated in 2012 in Northern Ireland seems to be completely different from the other VT2-phages from EAHEC, suggesting that at least two different VT-phages have been able to infect Enteroaggregative *E. coli* recipient strains.
The only trait in common between the two types of VT2-phages of EAHEC seems to be the presence of a gene encoding a tail fiber. Since it has been hypothesised that the infection with a lambdoid phage can be mediated by the cross-talking between the bacterium and the phage and tail fibers are part of this selection process (Werts *et al.*, 1994), we hypothesised that such a fiber might be part of the VT-phage host specificity machinery.

Final considerations
The studies conducted on the VT-phages as part of my PhD thesis, add experimental evidences to the hypothesis that the interplay between the capability of producing the Verocytotoxins and the ability of efficiently colonize the host intestinal mucosa is pivotal in the pathophysiology of VTEC infections. As a matter of fact, given the large variations in the VTEC types present in the reservoirs and the environment and the restricted number of them constantly isolated from human cases of disease, it can be hypothesised that for a VTEC to cause illness to man, particularly the most severe forms such as the HUS, an effective colonization of the host gut must intervene for the VTEC to win the competition with the endogenous microflora. It can also be stated that the stronger is the colonization, the harder is the offense caused to the host's intestinal barrier, resulting in a more severe disease. The thread VTEC to EHEC to EAHEC and the associated increasing severity of the disease caused is, in this sense, paradigmatic being the colonization of the host intestine operated by EAggEC the most extended and long lasting.
A further speculation derived from our studies concerns the finding that in VTEC O157 the phage Φ-8 seems to down-regulate the genes encoding the T3SS, providing a different scenario for the interplay between the toxin production and the colonization in causing the human illness. The pathogenic mechanism may have evolved so finely that the down-regulation of the LEE genes operated by Φ-8 could be a reflection of the need to tune up the transcription of a large number of virulence genes in response to the wide variations existing in a complex environment, such as the intestine is, in order to set the toxin's production at an optimal level in the moment the colonization has been successful.

Candidate's contributions to the present works:
Publication n. 1: Laura Grande contributed to the investigation of the phage's genome regions of interest in VTEC O157 strains by PCR, to the analysis of *vtx* gene transcription by Real-Time PCR and to the experiments aiming at the measurement of the LEE1 promoter activity. She helped in the drafting of the manuscript in the section "Materials and Methods" and revised critically the whole manuscript.
Publication n. 2: The candidate's contributions are described in a specific section in the respective paper.

Infection and Immunity

Identification and Characterization of a Peculiar *vtx2*-Converting Phage Frequently Present in Verocytotoxin-Producing Escherichia coli O157 Isolated from Human Infections

Rosangela Tozzoli, Laura Grande, Valeria Michelacci, Rosa Fioravanti, David Gally, Xuefang Xu, Roberto La Ragione, Muna Anjum, Guanghui Wu, Alfredo Caprioli and Stefano Morabito

Infect. Immun. **2014, 82(7):3023. DOI: 10.1128/IAI.01836-14. Published Ahead of Print 5 May 2014.**

IAI Journals.ASM.org

Identification and Characterization of a Peculiar *vtx2*-Converting Phage Frequently Present in Verocytotoxin-Producing *Escherichia coli* O157 Isolated from Human Infections

Rosangela Tozzoli,[a] Laura Grande,[a] Valeria Michelacci,[a] Rosa Fioravanti,[a]* David Gally,[b] Xuefang Xu,[b]* Roberto La Ragione,[c,d] Muna Anjum,[c] Guanghui Wu,[c] Alfredo Caprioli,[a] Stefano Morabito[a]

EU Reference Laboratory for *Escherichia coli*, Istituto Superiore di Sanità, Rome, Italy[a]; Immunity and Infection Division, Roslin Institute and R(D)SVS, University of Edinburgh, Easter Bush, Midlothian, United Kingdom[b]; Department of Bacteriology and Food Safety, Animal Health and Veterinary Laboratories Agency, Weybridge, Addlestone, Surrey, United Kingdom[c]; School of Veterinary Medicine, University of Surrey, Guildford, Surrey, United Kingdom[d]

Certain verocytotoxin-producing *Escherichia coli* (VTEC) O157 phage types (PTs), such as PT8 and PT2, are associated with severe human infections, while others, such as PT21, seem to be restricted to cattle. In an attempt to delve into the mechanisms underlying such a differential distribution of PTs, we performed microarray comparison of human PT8 and animal PT21 VTEC O157 isolates. The main differences observed were in the *vtx2*-converting phages, with the PT21 strains bearing a phage identical to that present in the reference strain EDL933, BP933W, and all the PT8 isolates displaying lack of hybridization in some regions of the phage genome. We focused on the region spanning the *gam* and *cII* genes and developed a PCR tool to investigate the presence of PT8-like phages in a panel of VTEC O157 strains belonging to different PTs and determined that a *vtx2* phage reacting with the primers deployed, which we named Φ8, was more frequent in VTEC O157 strains from human disease than in bovine strains. No differences were observed in the production of the VT2 mRNA when Φ8-positive strains were compared with VTEC O157 possessing BP933W. Nevertheless, we show that the *gam-cII* region of phage Φ8 might carry genetic determinants downregulating the transcription of the genes encoding the components of the type III secretion system borne on the locus of enterocyte effacement pathogenicity island.

Verocytotoxin (VT)-producing *Escherichia coli* (VTEC) O157 is a zoonotic pathogen causing food-borne disease outbreaks and sporadic cases of disease worldwide (1, 2). The symptoms induced upon VTEC O157 infection include a variety of clinical manifestations, such as diarrhea, hemorrhagic colitis, and the life-threatening hemolytic-uremic syndrome (HUS). VTEC O157 can be found as a component of the intestinal microflora in numerous animal species, but domestic ruminants, especially cattle, have been identified as its main reservoir (2). The presence of VTEC O157 in the intestinal content of cattle may cause the contamination of food of bovine origin at the slaughterhouse (3, 4). Moreover, healthy cattle shed VTEC O157 in their feces, contaminating the farm environment and favoring its persistence in the herd (5–7).

Although the main vehicle of infection is contaminated food of animal origin, the dispersion of VTEC O157 in the environment, caused by its elimination with ruminants' feces, also poses a risk for humans to acquire the infection. In particular, human infection can result from exposure to contaminated water, used either for drinking or for recreational purposes, as well as from consuming vegetables grown in fields irrigated with contaminated water or fertilized with animal manure not properly matured (8, 9).

The pathogenicity of VTEC O157 relies upon the expression of at least two key virulence features: the production of verocytotoxins (VTs), also termed Shiga toxins (Stxs), encoded by genes carried by temperate bacteriophages (10), and the induction of the characteristic attaching and effacing (A/E) lesion in the intestinal mucosa of the host (11), with the latter being conferred by the presence of a pathogenicity island termed the locus of enterocyte effacement (LEE) (12). The LEE harbors genes encoding several effectors involved in the pathogenesis of infections, such as an adhesin encoded by the gene *eae* and termed intimin; its translocated receptor, Tir; a type III secretion system (T3SS) (13); and a number of effectors delivered directly into the host cell via the T3SS and involved in the rearrangement of the enterocyte cytoskeleton.

Investigations of outbreaks caused by VTEC O157 are largely assisted by laboratory procedures aimed at subtyping the isolates, with the purposes of identifying the clusters of cases and tracing the vehicles of infection. Phage typing is one such typing technique that is able to distinguish about 80 phage types (PTs) according to the susceptibility of VTEC O157 to infection with a panel of bacteriophages (14). Although this technique was developed more than 2 decades ago, it still remains a useful approach to characterize VTEC O157 strains. Interestingly, it has been observed that while the isolates from cattle may span a wide portion of the entire PT panel (15), the strains isolated from both outbreaks and sporadic cases of human disease usually belong to a

Received 27 March 2014 **Returned for modification** 11 April 2014
Accepted 24 April 2014
Published ahead of print 5 May 2014
Editor: S. R. Blanke
Address correspondence to Rosangela Tozzoli, rosangela.tozzoli@iss.it.
* Present address: Rosa Fioravanti, Department of Biology, Università Roma TRE, Rome, Italy; Xuefang Xu, State Key Laboratory for Infectious Disease Prevention and Control (China CDC), National Institute of Communicable Diseases Control and Prevention and Control, Chinese Center for Disease Control and Prevention, Beijing, People's Republic of China.

doi:10.1128/IAI.01836-14

TABLE 1 Primer pairs designed in this study

Name	Sequence	Positions[a]	Size (bp) (accession no.)
Gam fwd	ATACCTCTGAATCAATATCAACCTG	1338126–1338150	6,106 (AE005174)
CII rev	AAAAGCACACAAGACCGAAG	1344231–1344212	
Roi2 fwd	GACAATGAATGAGCTGATAAATAGC	1349557–1349581	7,256 (AE005174)
S rev	ATATGTCAGCAGCCCAAACA	1356813–1356794	
Cro-CI up	AGAGCGGCTCCGCTTATTA	4685–4667	569 (KF241843)
Cro-CI low	TGAGTATTCGCCAACAGGTG	4116–4135	

[a] The positions refer to the accession numbers of the sequences used to deploy the primers (indicated in the size column).

restricted number of PTs. In particular, VTEC O157 strains isolated from human infections in Europe mainly belong to PT8 and PT2 (15–18), PT21/28 (19, 20), and PT32 (16). The uneven PT distribution between the strains isolated from human disease and the strains isolated from the animal reservoir seems to indicate that a subpopulation of VTEC O157 might have evolved that is either more virulent for the human host or better adapted to survive in the food chain. The existence of a distinct subpopulation of VTEC O157 has been demonstrated by molecular techniques (21, 22), including octamer-based genome scanning (23), single nucleotide polymorphism (SNP) analysis (24), and a lineage-specific polymorphism assay (25), supporting such a hypothesis.

In order to delve into the molecular bases underpinning this assumption, we carried out the comparative genomic analysis of VTEC O157 strains belonging to PT8, frequently isolated in Italy from cases of hemorrhagic colitis and HUS, and to PT21, which are commonly isolated from cattle but have been rarely associated with human cases.

In this paper, we show that the main genomic differences between the two groups of strains fell in the sequences of the bacteriophages carrying the *vtx2* genes and that the *vtx2*-converting phages present in most PT8 strains, whose prototype has been termed Φ8, are significantly more frequent among VTEC O157 strains from human infections than in bovine strains, regardless of their PTs. Moreover, we gathered indications that, in phage Φ8, one of these regions may carry genetic determinants downregulating the transcription of the LEE genes encoding components of the T3SS.

MATERIALS AND METHODS

Bacterial strains. The VTEC O157 strains isolated in Italy from different sources were part of the culture collection of the Reference Laboratory for *Escherichia coli* at the Istituto Superiore di Sanità. All the isolates possessed the intimin-coding *eae* gene (26) and produced VT, as assessed by a Vero cell cytotoxicity assay and PCR amplification of *vtx* genes (27). Phage typing was kindly performed at the Laboratory for Enteric Pathogens at Public Health England-Colindale, London, United Kingdom. The 20 strains belonging to PT8 and PT21 used in the microarray experiments were characterized for the presence of the enterohemolysin-coding gene and the 5′ fragment of the *efa1* gene by PCR amplification with primer pairs described previously (28, 29).

The VTEC O157 strains investigated for the presence of the *cro-cI* region of phage Φ8 included 138 Italian strains (100 of animal origin and 38 from human cases) and 30 PT21/28 bovine isolates from the culture collections held at the Roslin Institute (Edinburgh, United Kingdom).

The *E. coli* O157 strains EDL933 and RIMD0509952 Sakai and *E. coli* K-12 MG1655 were included in the study as reference strains. The *E. coli* K-12 strain JM109 was used in cloning experiments.

Microarray hybridizations. Microarray hybridizations and analysis were conducted at the Animal Health and Veterinary Laboratory Agency (AHVLA) in Weybridge, Surrey, United Kingdom. For each strain, data were compiled from two hybridizations. DNAs from each of the test strains were compared simultaneously, for gene presence or absence, to the whole genomes of the two *E. coli* O157 reference strains (EDL933 and RIMD0509952 Sakai) and the *E. coli* K-12 strain MG1655 on slides prepared in house containing about 6,000 oligonucleotides covering the complete open reading frames (ORFs) for the three control strains. Ten VTEC O157 strains belonging to PT8 and 10 strains belonging to PT21 were used in the microarray experiments. All the strains were isolated in Italy in the period 1993 to 2002. Total DNA was purified from each strain by using a genomic DNA extraction kit (Gentra Systems, USA) according to the manufacturer's instructions. Two micrograms of each test DNA was labeled using the BioPrime DNA-labeling system (InVitrogen Life Technologies, Carlsbad, CA, USA) with the Cy3 fluorophore, while a mixture at an equal concentration of the three control DNAs corresponding to a total of 2 μg was labeled with Cy5. The DNA was combined with 15 μg of random octamers, heated at 95°C for 5 min, and chilled on ice. The remaining components were added as follows: 0.12 mM dATP/GTP/TTP, 0.06 mM dCTP, and 0.01 mM Cy3- or Cy5-dCTP (final concentrations; GE Healthcare, Amersham, United Kingdom) and 40 units of the Klenow fragment of *E. coli* polymerase. The reaction mixture was placed at 37°C for 3 h, and the labeled DNA was purified using the Qiaquick PCR purification kit (Qiagen, Chatsworth, CA, USA) and eluted in 30 μl of water. Hybridizations were carried out for 16 to 18 h under glass coverslips in a sealed wet box at 65°C. Following hybridization, the slides were washed at room temperature for 2 min in two washing solutions (wash buffer 1, 1× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate]-0.05× SDS; wash buffer 2, 0.06× SSC) and dried by centrifugation in conical 50-ml tubes. The hybridized slides were scanned using a GenePix 4000B microarray scanner (Axon Instruments, Inc.), and the images were analyzed with BlueFuse Software (CB2 5LD; Bluegenome, Cambridge, United Kingdom). Finally, the results were analyzed with GeneSpring software (Agilent Technologies, CA, USA), which allows integration of the image data and their correlation with the list of target genes present in the reference strains.

Long and conventional PCR amplification of *vtx2*-converting phage regions. Long PCR amplifications were deployed to validate the microarray data and to investigate the presence and sizes of two regions of the *vtx2* bacteriophage spanning the *gam-cII* and *roi-s* genes, respectively. The primer pairs were designed on the sequence of the *vtx2* phage BP933W of the *E. coli* O157 reference strain EDL933 (GenBank accession no. AE005174). Primer sequences, together with their positions in the reference sequence and the expected sizes of the amplicons, are listed in Table 1. All PCRs were carried out using the TripleMaster PCR System (Eppendorf AG, Hamburg, Germany) and 200 ng of template DNA under the conditions indicated by the supplier.

A primer pair amplifying a DNA sequence internal to the *gam-cII* region (*cro-cI*) specific for the *vtx2* phage from the VTEC O157 PT8 strain ED257 (Table 1) was used to screen a wider collection of VTEC O157 isolates for the evaluation of the distribution of phages possessing such a region.

Determination of VTEC O157 lineages by LSPA-6 analysis. Lineage-specific polymorphism assay 6 (LSPA-6) was conducted using primers and multiplex PCR conditions described by Yang et al. (22). The primers

were labeled with 6-carboxyfluorescein (FAM) or hexachloroflorescein (Hex) and after amplification, the reactions were diluted 1:20 in distilled water. The fragments were separated by capillary electrophoresis using an Applied Biosystems 3130 Genetic Analyzer (Life Technologies, Grand Island, NY, USA) with a DS-30 matrix and carboxy-X-rhodamine (ROX)-labeled GeneFlo 625 (Chimerix, Milwaukee, WI, USA) as a size standard. Fragment sizes were assigned by using GeneMapper software v4.1 (Life Technologies, Grand Island, NY, USA), and the LSPA-6 alleles were determined on the basis of the respective reference sizes reported by Yang at al. (22). The isolates were grouped into lineages on the basis of the genotypes obtained according to the following definitions: strains possessing LSPA-6 genotype 111111 were classified as lineage I (LI) and isolates showing a 211111 profile as lineage I/II (LI/II), while all other allele combinations were classified as lineage II (LII) (30).

Cloning and sequencing of long PCR fragments. Long PCR fragments obtained for either the *gam-cII* or *roi-s* region were purified from the agarose gel with the Wizard SV Gel and PCR Clean-Up System (Promega Corporation, Madison, WI, USA) and cloned in pGEM-T Easy (Promega Corporation, Madison, WI, USA) under the conditions described in the user's manual supplied. Plasmid DNA was purified from *E. coli* K-12 JM109 using the FastPlasmid minikit (Eppendorf AG, Hamburg, Germany), and 1 μg was digested overnight with 20 U of EcoRI, NotI, and PstI restriction endonucleases to test the size of the cloned fragment.

Large-scale plasmid preparation was performed by using a Qiagen plasmid midikit (Qiagen, Chatsworth, CA, USA). The sequencing reactions were outsourced to the Sequencing Service Primm s.r.l., Milan, Italy. The DNA sequences obtained were compared with those present in the NCBI GenBank using the BLAST algorithm (31).

Bacteriophage induction and *vtx2* gene expression analysis. The 20 VTEC O157 strains used in the microarray experiments were grown to the exponential phase and treated with mitomycin C (0,5 μg/ml) to induce the bacteriophages. One milliliter of each bacterial culture was collected at different times after the phage induction: 0 min (not induced), 30 min, 1 h, 2 h, 3 h, and 4 h. Total RNA was extracted from 500 μl of bacterial cultures with an RNeasy minikit (Qiagen, Chatsworth, CA, USA).

DNase treatment of the RNA samples was done with the gDNA Wipeout 7× (Qiagen, Chatsworth, CA, USA), and cDNAs were prepared with QuantiTect reverse transcription (Qiagen, Chatsworth, CA, USA) using the conditions indicated by the suppliers. Ten nanograms of cDNA was used in real-time PCR experiments.

Primers and probes targeting the *vtx2* gene used in this study have been described previously (32). The *lacZ* gene real-time PCR amplification was conducted using primers and probes previously described (33) simultaneously with *vtx2* in order to normalize the fluorescence signals.

Analysis of T3SS-secreted proteins. Bacterial strains were cultured overnight in LB broth at 37°C with vigorous shaking. The cultures were diluted 1:100 in minimal essential medium (MEM)-HEPES (supplemented with 0.1% glucose, 25 mM sodium bicarbonate, and 0.25 μM ferric nitrate), grown to a final optical density at 600 nm (OD_{600}) of 0.5, and centrifuged at 4,000 × *g* for 15 min at 4°C. The supernatants were eventually filtered through 0.45-μm low-protein-binding filters (Millipore). The secreted proteins were precipitated using 10% (vol/vol) trichloroacetic acid (TCA) (Sigma-Aldrich) in the presence of bovine serum albumin (BSA) (4 μg/ml; New England BioLabs, United Kingdom) as the coprecipitant agent overnight at 4°C. Proteins were recovered by centrifugation at 4,000 × *g* for 30 min at 4°C. The protein pellets were air dried and dissolved in 1.5 M Tris-HCl, pH 8.8, buffer.

The secreted proteins were analyzed through SDS-12% PAGE and visualized by Coomassie blue staining or transferred onto a Hybond ECL nitrocellulose membrane (Amersham Biosciences) for Western blotting assays. The nitrocellulose membranes were saturated with 8% (wt/vol) skim milk powder (Oxoid) in phosphate-buffered saline (PBS) at 4°C overnight and incubated with anti-EspD monoclonal antibody (kindly provided by T. Chakraborty, University of Giessen, Giessen, Germany) diluted 1:5,000 in wash buffer (1% skim milk and 0.05% [vol/vol] polyoxyethylenesorbitan monolaurate [Tween 20] [Sigma-Aldrich] in PBS) and rabbit polyclonal anti-mouse IgG horseradish peroxidase (HRP)-conjugated antibodies (Jackson ImmunoResearch) diluted 1:500. The membranes were incubated for 2 h at room temperature (RT) on a platform shaker and washed three times for 10 min in wash buffer (1% skim milk and 0.05% Tween 20 in PBS) before and after each antibody step. For enhanced chemiluminescence (ECL) detection, membranes were incubated in 2.5 ml of ECL solution 1 mixed with 2.5 ml of ECL solution 2 (Amersham Biosciences, Glattbrugg, Switzerland) for 5 min at RT. Chemiluminescence was detected on Biomax-ML film (Kodak Industrie, Chalon sur Saon, France).

Measurement of LEE1 promoter activity. In order to evaluate the effect of the *gam-cII* region of Φ8 on the regulation of the genes present in the LEE1 operon, the plasmid pAJR71, containing a construct made up of a reporter gene encoding the green fluorescent protein (GFP) under the control of the LEE1 promoter, was used (34). The *E. coli* K-12 strain JM109 was cotransformed with the plasmids pAJR71 and pGEM-T Easy, where the 4.9-kb *gam-cII* region from Φ8 or the 6.1-kb region for BP933W were cloned. Control experiments were carried out, evaluating the production of GFP in the K-12 strain JM109 containing the pAJR71 plasmid, together with pGEM-T Easy without any insert. All the strains were cultured in Dulbecco's modified Eagle's medium supplemented with 15 μg/ml chloramphenicol at 37°C overnight. Subcultures were prepared by diluting (1:40) the overnight cultures in MEM-HEPES supplemented with 0.1% glucose, 25 mM sodium bicarbonate, and 0.25 μM ferric nitrate. Each subculture was grown to an OD_{600} of 0.5, and 200-μl aliquots were transferred into triplicate wells of a 96-well plate. The GFP produced by each subculture was assessed by reading the plate in a Victor 3 Multilabel Plate Reader (PerkinElmer, USA). The results were normalized by assessing the GFP production in at least three separate experimental sessions.

Nucleotide sequence accession number. The 4.9-kb *gam-cII* PCR fragment from the VTEC O157 PT8 strain ED257 sequence was submitted to GenBank under accession no. KF241843.

RESULTS

Microarray comparison of VTEC O157 strains belonging to PT8 and PT21. In order to investigate at the genomic level the differential distribution of PTs, we compared human VTEC O157 strains belonging to PT8 with strains belonging to PT21 of bovine origin by DNA-DNA microarray hybridization. Ten VTEC O157 strains belonging to PT8 and 10 belonging to PT21 were subjected to comparative genomics hybridization (CGH) experiments using microarray slides containing the whole complement of open reading frames from two VTEC O157 reference strains, EDL933 and RIMD0509952 Sakai, and from the *E. coli* K-12 strain MG1655. The main virulence traits of the investigated strains are reported in Table 2.

The CGH analysis showed that the PT8 and PT21 VTEC O157 strains investigated constituted two distinct clusters, with the exception of one PT8 (ED499) and one PT21 (ED350) strain (Fig. 1A). The analysis of the hybridization profiles showed that the main differences between the two groups of strains were in the DNA sequence of the *vtx2*-converting bacteriophage. In particular, the PT21 strains showed a complete pattern of hybridization with the ORFs corresponding to the *vtx2* bacteriophage BP933W of the reference strain EDL933, indicating the presence of a similar *vtx2* phage. On the other hand, the DNAs from PT8 strains did not hybridize with the BP933W ORFs in most of the phage genes, suggesting that, in these strains, the *vtx2* genes were located in a different type of bacteriophage. Moreover, when the cluster analysis was carried out considering the patterns of hybridization with

TABLE 2 Virulence characteristics of VTEC O157 strains selected for CGH experiments

Strain (phage type)	Source	Presence[a]				
		vtx1	*vtx2*	*eae*	*E-hly*	*Efa1-5′*
ED497 (PT8)	Human	+	+	+	+	+
ED507 (PT8)	Human	+	+	+	+	+
ED416 (PT8)	Human	+	+	+	+	+
ED421 (PT8)	Human	+	+	+	+	+
ED499 (PT8)	Human	−	+	+	+	+
ED307 (PT8)	Human	−	+	+	−	−
ED450 (PT8)	Human	+	+	+	+	+
ED472 (PT8)	Human	+	+	+	+	+
ED257 (PT8)	Human	+	+	+	+	+
ED159 (PT8)	Human	+	+	+	+	+
ED330 (PT21)	Cattle	−	+	+	−	−
ED438 (PT21)	Cattle	+	+	+	+	+
ED321 (PT21)	Cattle	+	+	+	+	+
ED207 (PT21)	Cattle	+	+	+	−	−
ED331 (PT21)	Cattle	+	+	+	−	−
ED314 (PT21)	Cattle	+	+	+	−	−
ED350 (PT21)	Cattle	−	+	+	−	−
ED326 (PT21)	Cattle	−	+	+	+	+
ED322 (PT21)	Cattle	+	+	+	+	+
ED281 (PT21)	Sheep	+	+	+	+	+

[a] +, present; −, absent.

the ORFs composing the *vtx2* bacteriophage only, a dendrogram coincident with that produced by considering the data from the entire genome was generated (Fig. 1B), confirming that the *vtx2* phage represented the major source of variability between the two groups of strains.

One of the polymorphic regions identified was that between the *gam* and *cII* genes, which is responsible for the switch between the lytic and the lysogenic cycles. The other polymorphic genes included those between *roi* and *s*, the late genes activated upon induction of the lytic cycle, including *vtx*. A few other regions, mainly containing phage structural genes, demonstrated absence of hybridization with respect to the BP933W sequence.

Investigation of the *gam-cII* and *roi-s* regions in PT8 and PT21 VTEC O157 strains. Two of the nonhybridizing regions of the *vtx2*-converting phage of PT8 VTEC O157 strains were further characterized by using a long-range PCR approach, and the results are reported in Table 3. All the PT8 strains tested in the microarray experiments produced an amplicon 4.9 kb in length when the entire *gam-cII* region was amplified, with the exception of strain ED307 (Table 3). Conversely, for the same region, all the PT21 strains produced a 6.1-kb amplification product, matching the predicted size of the *gam-cII* phage stretch present in phage BP933W. No amplification product was obtained for strain ED350 (Table 3).

As far as the *roi-s* region is concerned, 6 of the 10 PT8 strains produced an 8-kb amplicon, whereas an amplification product of 7.2 kb was obtained with 6 of the 10 PT21 strains. Again, the latter matched the expected size for the same region of the BP933W *vtx2* phage. Finally, in both groups, a few strains failed to yield amplification products or had varied product sizes, although different primer combinations designed on the same gene sequences were used (data not shown). This observation suggests that major polymorphisms in the sequences of the *roi* and *s* genes were also present (Table 3).

Characterization of the *gam-cII* region in the *vtx* phage of the ED257 strain. The region between *gam* and *cII* in lambda bacteriophages encodes several factors controlling the molecular switch between the lytic and lysogenic states of the phage, as well as other factors influencing the expression of late genes, which in the *vtx2*-converting phages include the genes encoding the verocytotoxins. Therefore, this region represented a good candidate for further work examining how it might influence the pathogenicity of VTEC O157 by affecting the level of *vtx* transcription.

The 4.9-kb *gam-cII* PCR fragment from one of the VTEC O157 PT8 strains (ED257) was cloned and sequenced. As expected, most of the DNA sequence showed low or no homology with the corresponding region on the BP933W phage. Conversely, a search among the sequences present in GenBank returned high homology (99%) with the sequence of the same region from the *vtx2*-converting phage of a VTEC O157 strain isolated during an outbreak that occurred in Japan in 1996 (35). This region was also similar to the DNA fragment comprising the *gam* and *cII* genes in the *vtx1*-converting phage CP933V in the VTEC O157 reference strain EDL933 (GenBank accession no. AE005174). To distinguish the *vtx2* phage identified in the PT8 strain ED257, which we termed Φ8, from the CP933V-like *vtx1* phage, we designed a PCR primer pair able to specifically amplify the 569-bp region between the *cro* and *cI* genes in the sequence of the *vtx2* phage of strain ED257. This PCR was used to assess the presence of phages possessing the *cro-cI* region of Φ8 in a panel of VTEC O157 strains isolated in Italy and including 38 strains from human infections and 100 strains of animal origin. The isolates belonged to different PTs and displayed different *vtx* gene profiles. The results of this PCR screening (Table 4) showed the presence of a Φ8-specific *cro-cI* region in 81.6% of the human isolates, whereas only 60% of the animal strains were positive in the assay ($P < 0{,}001$). These results suggest that the presence of Φ8-like phages is predominant among the VTEC O157 strains causing human infections, regardless of the PTs they belong to.

The Φ8-specific PCR assay was also used to analyze 30 VTEC O157 strains isolated in the United Kingdom and belonging to PT21/28, the PTs most frequently observed among the strains isolated from human infections in that country (3, 20). In agreement with the high frequency observed among the Italian human isolates, most of the PT21/28 strains investigated (28 out of 30) were positive in the *cro-cI* PCR. It is noteworthy that PT21/28 VTEC O157 strains have been associated with high excretion levels from cattle (20).

Characterization of LSPA-6 genotypes of the Italian VTEC O157. The LSPA-6 analysis of 138 VTEC O157 strains isolated in Italy showed that 66 out of the 138 Italian VTEC O157 strains (47.8%) belonged to the LI/II lineage. Interestingly, this represents an intermediate rate compared with the reported frequencies for this VTEC O157 lineage: 85% for the Australian strains (30), 16% for the VTEC O157 strains isolated in the United States (30), and 90% of the isolates from Argentina (36). Similarly, the distribution of the LI lineage among the Italian isolates was 16% versus 2% reported in a similar study involving VTEC O157 strains from Australia (30), 60% reported for the United States (30), and 4% reported for Argentina (36). The LII lineage had a higher prevalence in Italy, with 36,2% of the isolates tested, while it could be assigned to 13% of Australian isolates and 25% of the VTEC O157 strains from the United States (30). As for the relative distribution of the lineages, most of the human isolates belonged

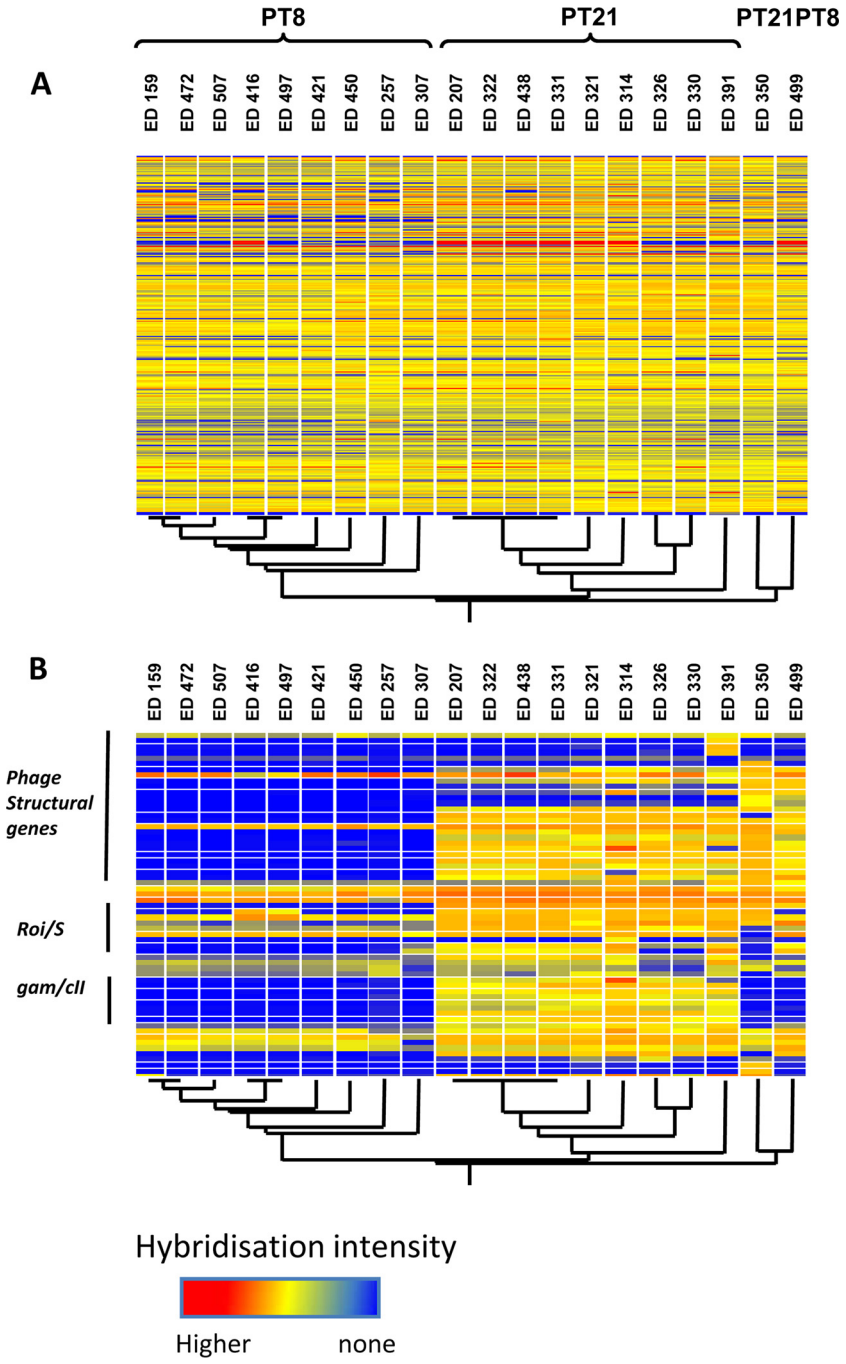


FIG 1 CGH analysis of microarray hybridization profiles of VTEC O157 strains. (A) Clustering of the analyzed strains produced by considering the whole ORF content. (B) Clustering of the strains obtained using only the ORFs composing the *vtx2*-converting bacteriophage. The dendrograms were constructed on the basis of the different intensities of the hybridization signals (from blue, negative, to red, high hybridization signal) by means of Genespring software (Agilent Technologies), using the Pearson correlation.

to LI/II (65%), followed by LII (26%) and LI (9%). The VTEC O157 strains isolated from animal sources belonged to similar proportions of the LI/II and LII lineages (42% and 40%, respectively), with only 18% of the isolates from LI. LSPA-6 typing showed that 91% of the human isolates belonged to lineages LI/II and LII, 85% of which were also positive in the *cro-cI* PCR. Interestingly, only the 68% of animal VTEC O157 isolates belonging to lineages LI/II and LII possessed a *vtx2* phage with the *cro-cI* region of Φ8.

Analysis of *vtx* gene transcription. The observed association of Φ8-like phages with VTEC O157 strains from human infections prompted us to investigate further. Since in the *vtx*-converting bacteriophages the *vtx* genes are under the control of the late gene promoter (37), the possibility that the *vtx2* genes carried by Φ8-

TABLE 3 Long-range PCR analysis of the *gam-cII* and *roi-s* regions of VTEC O157 strains included in the microarray experiments

Strain	Amplicon size (kb)[a]	
	gam-cII	*roi-s*
ED497 (PT8)	4.9	Neg
ED507 (PT8)	4.9	8
ED416 (PT8)	4.9	8
ED421 (PT8)	4.9	Neg
ED499 (PT8)	4.9	8
ED307 (PT8)	8 + 3	Neg
ED450 (PT8)	4.9	8
ED472 (PT8)	4.9	Neg
ED257 (PT8)	4.9	8
ED159 (PT8)	4.9	8
ED330 (PT21)	6.1	Neg
ED438 (PT21)	6.1	7.2
ED321 (PT21)	6.1	Neg
ED207 (PT21)	6.1	7.2
ED331 (PT21)	6.1	7.2
ED314 (PT21)	6.1	7.2
ED350 (PT21)	Neg	Neg
ED326 (PT21)	6.1	7.2
ED322 (PT21)	6.1 + 4.9	7.2
ED281 (PT21)	6.1 + 4.9	6

[a] Neg, negative.

like phages might produce increased levels of VT mRNA was investigated. The transcription of such genes is boosted upon induction of the lytic cycle, when the gene *N*, present in the *gam-cII* region and encoding an antiterminator, is activated, allowing the transcription to proceed through a terminator site. This event triggers the transcription of another antiterminator, the product of the gene *Q*, which in turn allows the transcript to run over another termination site located upstream of the *vtx* genes (37). Therefore, we investigated the possibility that VTEC O157 possessing Φ8-like phages produced higher levels of VT mRNA than those with a BP933W-like *vtx2* phage. Four VTEC O157 PT8 strains possessing *vtx2* phages with the *gam-cII* region of Φ8 and four PT21 strains harboring a single BP933W-like *vtx2* phage were included in the experiment. The amount of *vtx2A* mRNA was measured by reverse transcriptase PCR at different intervals after inducing the *vtx2* phage by the addition of mitomycin C. The results of the assays showed that the amounts of *vtx2A* mRNA increased as the induction progressed, but no significant differences between the two groups of strains were observed (Table 5), indicating that, at least under laboratory conditions, the presence of Φ8-like phages does not enhance the production of the *vtx2A* mRNA.

Influence of the Φ8 *gam-cII* region on T3SS production. Recently, it has been shown that factors encoded on prophages can influence the regulation of the LEE (38, 39), which governs the induction of A/E lesions via the production and assembly of a complete T3SS. It has also been proposed that variations in the expression of the T3SS by *E. coli* O157 strains could have an impact on the colonization of the host (40). These assumptions prompted us to investigate if the presence of the Φ8 phage might influence the transcription of LEE genes in VTEC O157. The expression of T3SS components was evaluated by assessing the amount of EspD protein produced and secreted. EspD is part of, and is also secreted through, the T3SS machinery, thus represent-

TABLE 4 Identification of Φ8-like phages in Italian VTEC O157 strains belonging to different PTs and isolated from different sources by PCR amplification of the *cro-cI* region

Source	PT	No. of strains	Presence[a]		
			vtx1	*vtx2*	*cro-cI*
Human[b]	1	1	−	+	+
	2	3	−	+	−
	2	5	−	+	+
	4	1	+	+	−
	4	1	−	+	+
	8	2	−	+	+
	8	9	+	+	+
	14	2	−	+	+
	14	3	+	+	+
	20	1	−	+	+
	21	1	−	+	−
	32	1	−	+	+
	34	1	−	+	+
	43	1	−	+	+
	49	3	−	+	+
	54	1	−	+	−
	56	1	−	+	−
	21/28	1	+	+	+
Animals and foodsuffs[c]	1	2	+	+	−
	1	1	+	+	+
	1	2	−	+	+
	2	2	−	+	+
	2	11	−	+	−
	3	1	−	+	+
	4	1	+	+	−
	4	1	+	+	+
	4	2	−	+	+
	8	1	+	+	−
	8	1	−	+	−
	8	1	−	+	+
	8	12	+	+	+
	14	10	−	+	+
	20	2	−	+	−
	20	2	−	+	+
	21	3	−	+	−
	21	6	+	+	−
	23	2	+	+	+
	31	3	−	+	−
	31	4	−	+	+
	32	1	+	−	−
	32	1	−	+	+
	32	1	+	+	−
	33	1	−	+	−
	33	2	+	+	−
	34	1	−	+	−
	34	7	−	+	+
	43	1	−	+	−
	43	1	−	+	+
	43	1	+	+	+
	44	1	−	+	+
	49	2	−	+	+
	51	1	−	+	−
	54	5	−	+	+
	54	1	+	+	+
	63	1	−	+	+
	21/28	2	−	+	−

[a] +, present; −, absent.
[b] Total no. of strains, 38; no. *cro-cI* positive, 31 (81.6%) ($P < 0.025$).
[c] Total no. of strains, 100; no. *cro-cI* positive, 60 (60%).

TABLE 5 Evaluation of *vtx2* gene expression after induction with mitomycin C

Strain	*vtx2* phage	$C_{T\ value}$[a]					
		Time zero	0.5 h	1 h	2 h	3 h	4 h
ED220	BP933W	21	25.9	27	19.5	16.4	13.6
ED419	BP933W	24.5	20.5	20	12.7	16	9
ED250	BP933W	23	22.5	22	15.2	13.3	11.8
ED320	BP933W	22.8	20.3	18	14	10	16
ED499	Φ8-like	25	22.5	21.3	15	17	17
ED417	Φ8-like	17	21	20	13.8	15	15
ED154	Φ8-like	25	26	22	17	14.3	14.3
ED254	Φ8-like	0	28	23.8	21	19	19

[a] For each strain, the C_T values of the real-time PCR amplification of the *vtx2* cDNA are reported at different times after induction.

ing a good marker to evaluate the level of T3SS component expression. Ten strains belonging to PT8 and 10 strains belonging to PT21, possessing the Φ8-like phages or the BP933W-like phage, respectively, were examined. Western blot analyses indicated a marked difference in the relative amounts of EspD secreted in the culture supernatants, with higher levels produced by VTEC O157 strains belonging to PT21 than by those in PT8 (Fig. 2A). This finding was in agreement with the previous observation that VTEC O157 strains belonging to PT21/28, which also harbor *vtx2* phages resembling Φ8 at high frequency, secreted significantly less EspD than VTEC O157 belonging to PT32, which was used as a comparative group (38).

To evaluate if the presence of the *gam-cII* region of phage Φ8 could directly influence the production of EspD, we studied the effect of the cloned *gam-cII* region of Φ8 on the expression of the GFP gene cloned under the control of the promoter regulating the transcription of the LEE1 operon on the LEE in comparison with that exerted by the same region cloned from BP933W in an *E. coli* K-12 background. Since the LEE1 operon encodes the structural components of the T3SS, the system can provide information about the influence of these phage regions, if any, on the T3SS regulation mechanisms through the analysis of the level of green fluorescent protein produced. Such an experimental model showed that the 4.9-kb *gam-cII* region from Φ8 induced dramatic repression of the LEE1 promoter ($P < 0.0001$; Mann-Whitney test) compared to the effect observed when the same region from BP933W was cotransformed in strain K-12 with the LEE1-GFP construct (Fig. 2B). The latter combination did not show significant differences from the same system containing the construct LEE1-GFP in the presence of the plasmid used for cloning the phage regions but containing no inserts. This result suggests that one or more factors, encoded by genes present in the *gam-cII* region of the Φ8 phage, may negatively influence the transcription of the genes under the control of the LEE1 promoter.

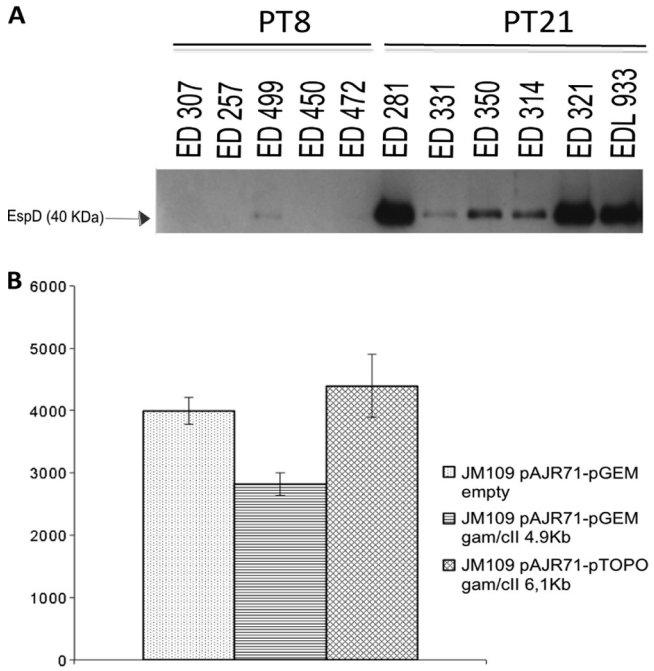


FIG 2 Analysis of T3SS protein production. (A) Western blot analysis of EspD production in VTEC O157 strains belonging to PT8 and PT21. EDL933 was included as a control. (B) GFP assay showing LEE1 promoter activity in the *E. coli* K-12 strain JM109 cotransformed with either pAJR71 and pGEM-gam/cII-Φ8 or pAJR71 and pGEM-gam/cII-BP933W. The first bar corresponds to the GFP levels detected in the *E. coli* K-12 strain JM109 cotransformed with pAJR71 and the pGEM-T Easy empty vector. The error bars indicate standard deviations.

DISCUSSION

The hypothesis that different VTEC O157 clones could be characterized by higher virulence or more efficiently transmitted to the human host has been formulated by several authors, based on molecular characterization studies showing that strains isolated from cattle and from human cases of disease often belong to different clusters (21, 23, 41, 42). Accordingly, we have observed that the VTEC O157 strains isolated in Italy from human cases of infection and from animal sources are differentially distributed within the lineages identified by the LSPA-6 assay, with the majority of the human isolates belonging to the lineages LI/II and LII. Interestingly, the Italian VTEC O157 strains were positioned differently from the isolates reported from the United States (30), Australia (30), and Argentina (36), suggesting a geographically driven clonal development.

The existence of a VTEC O157 subpopulation has also been supported by the observation that VTEC O157 strains isolated from cases of human disease usually belong to a nonrandom subset of PTs (15–20). Although the phage types are related to the susceptibility of VTEC O157 to infection with a panel of phages and may not correlate with their virulence potentials, significant differences have been reported in the distribution of the PTs among VTEC O157 isolates from human and bovine sources by different authors (15–20) and can thus be considered a good epidemiological marker for the purpose of identifying VTEC O157

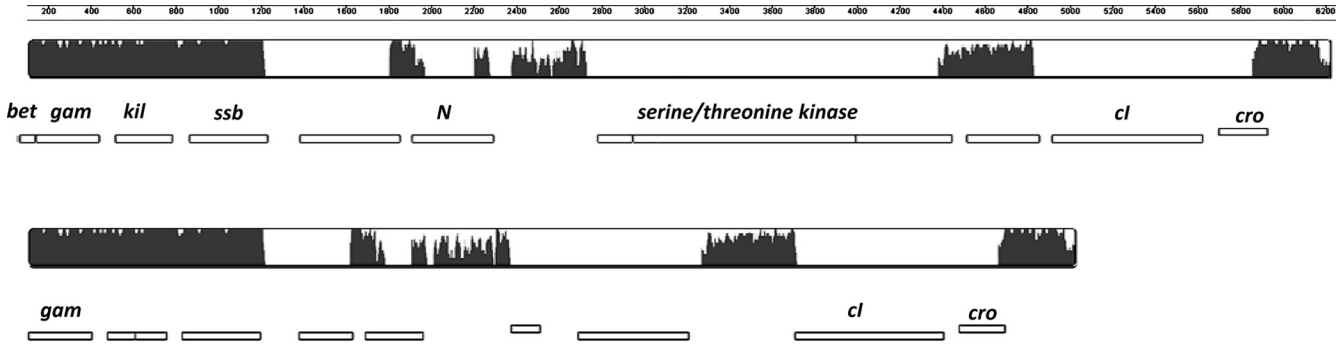


FIG 3 Alignment of the nucleotide sequences of the *gam-cII* region present in the *vtx2* phage from the strains EDL933 (top) and ED257 (bottom) obtained with progressive MAUVE alignment software (47). Only the names of known genes are reported. All the unnamed genes (white boxes) correspond to hypothetical protein-coding genes. The regions showing sequence similarities in the two samples are colored in gray, with the height of the colored areas corresponding to the alignment score at each position.

subpopulations. Therefore, we based our investigation on the observation that, in Italy, about half of human infections with VTEC O157 are caused by strains of PT8, while only 1 case of infection out of the 45 cases microbiologically confirmed and reported to the Italian HUS registry in the period 1988 to 2006 was caused by a VTEC O157 strain of PT21. In order to investigate the genetic differences underlying such an uneven distribution of PTs, we carried out a comparative genomic analysis of VTEC O157 strains belonging to PT8 and PT21, which, despite its low frequency in the human isolates, is common among bovine strains.

This analysis led to the identification of a region present in the *vtx2*-converting phages of the human PT8 strains, whose prototype has been termed Φ8 in the VTEC O157 PT8 strain ED257, that significantly differed from its homologous region in phage BP933W, the *vtx2*-converting phage present in the VTEC O157 reference strain EDL933 and in the bovine PT21 Italian strains. Such a region, including the genes between *gam* and *cII* (Fig. 3), regulates the switch between the lytic and lysogenic cycles.

Since the induction of prophages carrying the *vtx* genes is a key event in the regulation of the *vtx* genes and boosts production of the VT mRNA, the polymorphism detected in this region was further investigated. Sequencing of the *gam-cII* fragment of phage Φ8 showed similarity with the sequence of the same region of the *vtx2*-converting phage harbored by the VTEC O157 strain Morioka V526, which caused a large outbreak of infections in Japan during the 1990s (35). By using a primer pair specifically targeting the *cro-cI* region of phage Φ8, we observed that this region is present in *vtx2* phages from the large majority of VTEC O157 strains from human infections isolated in Italy regardless of their PTs (Table 4), indicating that this peculiar phage region is not a marker for PT8 strains but rather identifies a *vtx2* phage, or a family of *vtx2* phages, segregating with VTEC O157 that causes disease in humans. We also observed that almost all (28 out of 30) of the VTEC O157 strains belonging to PT21/28 isolated in the United Kingdom and assayed in this study harbored a *vtx2*-converting phage possessing the *cro-cI* region of Φ8, while only two of them possessed a different type of phage that was negative in the *cro-cI* PCR. Interestingly, VTEC O157 PT21/28 strains are commonly isolated from cases of severe human disease in the United Kingdom (19) and have also been associated with the supershedding phenotype in cattle (4). Given the strong association of Φ8-like phages with strains isolated from human illness, we hypothesized that their presence might favor the induction of disease by triggering the production of larger amounts of VT2 than the *vtx2* phages commonly found in VTEC O157 strains populating the animal reservoir. However, quantification of VT2 mRNA did not show differences in the levels of expression between strains harboring *vtx2* phages similar to Φ8 and strains harboring BP933W-like phages. This observation suggests that the presence of a *vtx2* phage possessing the *cro-cI* region of phage Φ8 does not influence the virulence potential of VTEC O157 by inducing augmented levels of VTs. Therefore, the observed association of this phage with strains from human disease must have a different cause.

Besides the production of VTs, the ability to colonize the intestinal mucosa by inducing a T3SS-mediated attaching and effacing lesion is considered to be pivotal to the pathophysiology of VTEC O157-induced disease (43). The T3SS is assembled from a number of components produced by genes harbored by the LEE pathogenicity island, which also includes genes encoding the adhesin intimin and its T3SS-translocated receptor, Tir (44, 45). The production and assembly of the T3SS is finely regulated by several factors encoded by genes present either in the LEE itself or on other genomic structures, such as the one encoding the prophage regulator RgdR, located on the phage-derived O island 51 in the reference strain EDL933, whose effect on the expression of the T3SS components has been highlighted (46).

In this respect, the association of Φ8-like phages with PT21/28 VTEC O157 strains described here is noteworthy. PT21/28 VTEC strains have been shown to produce levels of EspD, an effector translocated via the T3SS, lower than those produced by the VTEC O157 reference strain EDL933 (38). Accordingly, Italian Φ8-positive PT8 strains produced smaller amounts of EspD than the PT21 strains that harbored BP933W-like phages (Fig. 2A).

Given these observations, we explored the possibility that the presence of Φ8-like phages may also control T3SS expression. Since the genes encoding the T3SS components are under the control of the LEE1 promoter in the LEE pathogenicity island (46), we measured the production of GFP by an *E. coli* K-12 strain containing a GFP-coding gene cloned downstream of the LEE1 promoter in the presence of the *gam-cII* region from Φ8 or from the BP933W phages. This approach clearly showed that the presence of the DNA region from the Φ8 phage inhibited the production of GFP while the corresponding BP933W region did not have an effect on the LEE1-controlled transcription of GFP. These re-

sults suggest that the *gam-cII* region of phage Φ8 contains one or more regulators influencing, directly or indirectly, the transcription of the LEE1 promoter and, consequently, T3SS expression. This finding correlates with reduced production of EspD observed in PT8 strains (Fig. 2A) and PT21/28 VTEC O157 strains, as previously described (38).

EspD is part of the T3SS translocation apparatus and is required for effector delivery into host cells (46). Therefore, altered levels of EspD are likely to have an impact on epithelial cell colonization. The differences observed in EspD production and their correlation with the presence of Φ8-like phages indicate that these phages are likely to influence VTEC O157 colonization of the gastrointestinal tract of the host.

Although the biological significance of the observed effects of the Φ8 *gam-cII* region on LEE1 transcription still needs to be elucidated, the high frequency of *vtx2* phages displaying the presence of such a region in PT21/28 VTEC O157 strains may be of help in understanding the association of Φ8-like phages with VTEC O157 isolates from human infections. In fact PT21/28 strains are frequently isolated from supershedding cattle (4), and the finding that Φ8-like phages are common in these isolates suggests that the fine tuning of T3SS expression may play a role in establishing the supershedding status. In turn, since supershedding is important for VTEC O157 to be established and maintained in the herd and is a critical risk factor for human infections (20), the regulation of the T3SS exerted by Φ8-like *vtx2* phages may cause increased exposure of humans to VTEC O157 through supershedding, eventually explaining the observed overrepresentation of Φ8 in human strains.

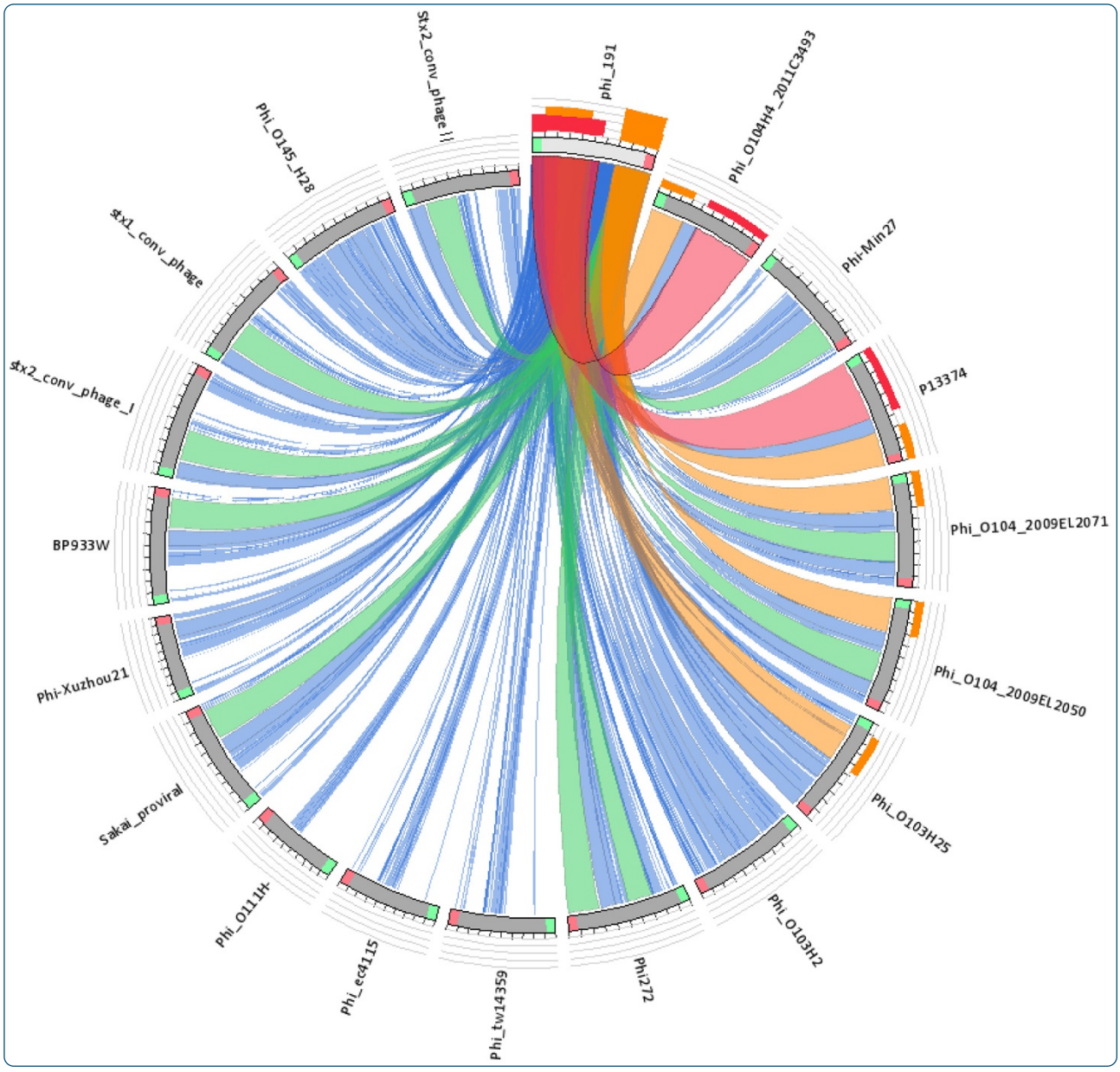
Further studies are needed to ascertain how the presence of the Φ8 phage in VTEC O157 contributes to colonization of the gastrointestinal tract.

REFERENCES

1. **Griffin PM, Tauxe RV.** 1991. The epidemiology of infections caused by Escherichia coli O157:H7, other enterohemorrhagic E. coli, and the associated hemolytic uremic syndrome. Epidemiol. Rev. **13:**60–98.
2. **Armstrong GL, Hollingsworth J, Morris JG, Jr.** 1996. Emerging foodborne pathogens: Escherichia coli O157:H7 as a model of entry of a new pathogen into the food supply of the developed world. Epidemiol Rev. **18:**29–51. http://dx.doi.org/10.1093/oxfordjournals.epirev.a017914.
3. **Chase-Topping ME, McKendrick IJ, Pearce MC, MacDonald P, Matthews L, Halliday J, Allison L, Fenlon D, Low JC, Gunn G, Woolhouse ME.** 2007. Risk factors for the presence of high-level shedders of Escherichia coli O157 on Scottish farms. J. Clin. Microbiol. **45:**1594–1603. http://dx.doi.org/10.1128/JCM.01690-06.
4. **Chase-Topping M, Gally D, Low C, Matthews L, Woolhouse M.** 2008. Super-shedding and the link between human infection and livestock carriage of Escherichia coli O157. Nat. Rev. Microbiol. **6:**904–912. http://dx.doi.org/10.1038/nrmicro2029.
5. **Michel P, Wilson JB, Martin SW, Clarke RC, McEwen SA, Gyles CL.** 1999. Temporal and geographical distributions of reported cases of Escherichia coli O157:H7 infection in Ontario. Epidemiol. Infect. **122:**193–200. http://dx.doi.org/10.1017/S0950268899002083.
6. **Low JC, McKendrick IJ, McKechnie C, Fenlon D, Naylor SW, Currie C, Smith DG, Allison L, Gally DL.** 2005. Rectal carriage of enterohemorrhagic Escherichia coli O157 in slaughtered cattle. Appl. Environ. Microbiol. **71:**93–97. http://dx.doi.org/10.1128/AEM.71.1.93-97.2005.
7. **Hancock DD, Besser TE, Kinsel ML, Tarr PI, Rice DH, Paros MG.** 1994. The prevalence of Escherichia coli O157.H7 in dairy and beef cattle in Washington State. Epidemiol. Infect. **113:**199–207. http://dx.doi.org/10.1017/S0950268800051633.
8. **O'Brien SJ, Adak GK, Gilham C.** 2001. Contact with farming environment as a major risk factor for Shiga toxin (Vero cytotoxin)-producing Escherichia coli O157 infection in humans. Emerg. Infect. Dis. **7:**1049–1051. http://dx.doi.org/10.3201/eid0706.010626.
9. **Strachan NJ, Dunn GM, Locking ME, Reid TM, Ogden ID.** 2006. Escherichia coli O157: burger bug or environmental pathogen? Int. J. Food Microbiol. **112:**129–137. http://dx.doi.org/10.1016/j.ijfoodmicro.2006.06.021.
10. **Huang A, Friesen J, Brunton JL.** 1987. Characterization of a bacteriophage that carries the genes for production of Shiga-like toxin 1 in Escherichia coli. J. Bacteriol. **169:**4308–4312.
11. **Frankel G, Phillips AD, Rosenshine I, Dougan G, Kaper JB, Knutton S.** 1998. Enteropathogenic and enterohaemorrhagic Escherichia coli: more subversive elements. Mol. Microbiol. **30:**911–921. http://dx.doi.org/10.1046/j.1365-2958.1998.01144.x.
12. **McDaniel TK, Jarvis KG, Donnenberg MS, Kaper JB.** 1995. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. Proc. Natl. Acad. Sci. U. S. A. **92:**1664–1668. http://dx.doi.org/10.1073/pnas.92.5.1664.
13. **Delahay RM, Frankel G, Knutton S.** 2001. Intimate interactions of enteropathogenic Escherichia coli at the host cell surface. Curr. Opin. Infect. Dis. **14:**559–565. http://dx.doi.org/10.1097/00001432-200110000-00009.
14. **Khakhria R, Duck D, Lior H.** 1990. Extended phage-typing scheme for Escherichia coli O157:H7. Epidemiol. Infect. **105:**511–520. http://dx.doi.org/10.1017/S0950268800048135.
15. **Mora A, Blanco M, Blanco JE, Alonso MP, Dhabi G, Thomson-Carter F, Usera MA, Bartolome R, Prats G, Blanco J.** 2004. Phage types and genotypes of Shiga toxin-producing Escherichia coli O157:H7 isolates from humans and animals in Spain: identification and characterization of two predominating phage types (PT2 and PT8). J. Clin. Microbiol. **42:**4007–4015. http://dx.doi.org/10.1128/JCM.42.9.4007-4015.2004.
16. **Kappeli U, Hachler H, Giezendanner N, Cheasty T, Stephan R.** 2011. Shiga toxin-producing Escherichia coli O157 associated with human infections in Switzerland, 2000–2009. Epidemiol. Infect. **139:**1097–1104. http://dx.doi.org/10.1017/S0950268810002190.
17. **Stirling A, McCartney G, Ahmed S, Cowden J.** 2007. An outbreak of Escherichia coli O157 phage type 2 infection in Paisley, Scotland. Euro Surveill. **12:**E070823.1.
18. **Hart J, Smith G.** 2009. Verocytotoxin-producing Escherichia coli O157 outbreak in Wrexham, North Wales, July 2009. Euro Surveill. **14:**19300.
19. **Locking M, Allison L, Rae L, Pollock K, Hanson M.** 2006. VTEC infections and livestock-related exposures in Scotland, 2004. Euro Surveill. **11:**E060223.4.
20. **Matthews L, Reeve R, Gally DL, Low JC, Woolhouse ME, McAteer SP, Locking ME, Chase-Topping ME, Haydon DT, Allison LJ, Hanson MF, Gunn GJ, Reid SW.** 2013. Predicting the public health benefit of vaccinating cattle against Escherichia coli O157. Proc. Natl. Acad. Sci. U. S. A. **110:**16265–16270. http://dx.doi.org/10.1073/pnas.1304978110.
21. **Dowd SE, Ishizaki H.** 2006. Microarray based comparison of two Escherichia coli O157:H7 lineages. BMC Microbiol. **6:**30. http://dx.doi.org/10.1186/1471-2180-6-30.
22. **Yang Z, Kovar J, Kim J, Nietfeldt J, Smith DR, Moxley RA, Olson ME, Fey PD, Benson AK.** 2004. Identification of common subpopulations of non-sorbitol-fermenting, beta-glucuronidase-negative Escherichia coli O157:H7 from bovine production environments and human clinical samples. Appl. Environ. Microbiol. **70:**6846–6854. http://dx.doi.org/10.1128/AEM.70.11.6846-6854.2004.
23. **Kim J, Nietfeldt J, Benson AK.** 1999. Octamer-based genome scanning distinguishes a unique subpopulation of Escherichia coli O157:H7 strains in cattle. Proc. Natl. Acad. Sci. U. S. A. **96:**13288–13293. http://dx.doi.org/10.1073/pnas.96.23.13288.
24. **Manning SD, Motiwala AS, Springman AC, Qi W, Lacher DW, Ouellette LM, Mladonicky JM, Somsel P, Rudrik JT, Dietrich SE, Zhang W, Swaminathan B, Alland D, Whittam TS.** 2008. Variation in virulence among clades of Escherichia coli O157:H7 associated with disease outbreaks. Proc. Natl. Acad. Sci. U. S. A. **105:**4868–4873. http://dx.doi.org/10.1073/pnas.0710834105.
25. **van Hoek AH, Aarts HJ, Bouw E, van Overbeek WM, Franz E.** 2013. The role of rpoS in Escherichia coli O157 manure-amended soil survival and distribution of allelic variations among bovine, food and clinical isolates. FEMS Microbiol. Lett. **338:**18–23. http://dx.doi.org/10.1111/1574-6968.12024.
26. **Oswald E, Schmidt H, Morabito S, Karch H, Marches O, Caprioli A.** 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. http://dx.doi.org/10.1128/IAI.68.1.64-71.2000.
27. **Morabito S, Karch H, Schmidt H, Minelli F, Mariani-Kurkdjian P, Allerberger F, Bettelheim KA, Caprioli A.** 1999. Molecular characterisation of verocytotoxin-producing Escherichia coli of serogroup O111 from different countries. J. Med. Microbiol. **48:**891–896. http://dx.doi.org/10.1099/00222615-48-10-891.
28. **Schmidt H, Beutin L, Karch H.** 1995. Molecular analysis of the plasmid-encoded hemolysin of Escherichia coli O157:H7 strain EDL 933. Infect. Immun. **63:**1055–1061.
29. **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:**3343–3348. http://dx.doi.org/10.1128/IAI.71.6.3343-3348.2003.
30. **Mellor GE, Besser TE, Davis MA, Beavis B, Jung W, Smith HV, Jennison AV, Doyle CJ, Chandry PS, Gobius KS, Fegan N.** 2013. Multilocus genotype analysis of Escherichia coli O157 isolates from Australia and the United States provides evidence of geographic divergence. Appl. Environ. Microbiol. **79:**5050–5058. http://dx.doi.org/10.1128/AEM.01525-13.
31. **Karlin S, Altschul SF.** 1990. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. Proc. Natl. Acad. Sci. U. S. A. **87:**2264–2268. http://dx.doi.org/10.1073/pnas.87.6.2264.
32. **Perelle S, Dilasser F, Grout J, Fach P.** 2004. Detection by 5′-nuclease PCR of Shiga-toxin producing Escherichia coli O26, O55, O91, O103, O111, O113, O145 and O157:H7, associated with the world's most frequent clinical cases. Mol. Cell. Probes **18:**185–192. http://dx.doi.org/10.1016/j.mcp.2003.12.004.
33. **Ram S, Shanker R.** 2005. Computing TaqMan probes for multiplex PCR detection of E. coli O157 serotypes in water. In Silico Biol. **5:**499–504.
34. **Roe AJ, Yull H, Naylor SW, Woodward MJ, Smith DG, Gally DL.** 2003. Heterogeneous surface expression of EspA translocon filaments by Escherichia coli O157:H7 is controlled at the posttranscriptional level. Infect. Immun. **71:**5900–5909. http://dx.doi.org/10.1128/IAI.71.10.5900-5909.2003.
35. **Watarai M, Sato T, Kobayashi M, Shimizu T, Yamasaki S, Tobe T, Sasakawa C, Takeda Y.** 1998. Identification and characterization of a newly isolated Shiga toxin 2-converting phage from Shiga toxin-producing Escherichia coli. Infect. Immun. **66:**4100–4107.
36. **Mellor GE, Sim EM, Barlow RS, D'Astek BA, Galli L, Chinen I, Rivas M, Gobius KS.** 2012. Phylogenetically related Argentinean and Australian Escherichia coli O157 isolates are distinguished by virulence clades and alternative Shiga toxin 1 and 2 prophages. Appl. Environ. Microbiol. **78:**4724–4731. http://dx.doi.org/10.1128/AEM.00365-12.
37. **Wagner PL, Neely MN, Zhang X, Acheson DW, Waldor MK, Friedman DI.** 2001. Role for a phage promoter in Shiga toxin 2 expression from a pathogenic Escherichia coli strain. J. Bacteriol. **183:**2081–2085. http://dx.doi.org/10.1128/JB.183.6.2081-2085.2001.
38. **Xu X, McAteer SP, Tree JJ, Shaw DJ, Wolfson EB, Beatson SA, Roe AJ, Allison LJ, Chase-Topping ME, Mahajan A, Tozzoli R, Woolhouse ME, Morabito S, Gally DL.** 2012. Lysogeny with Shiga toxin 2-encoding bacteriophages represses type III secretion in enterohemorrhagic Escherichia coli. PLoS Pathog. **8:**e1002672. http://dx.doi.org/10.1371/journal.ppat.1002672.
39. **Tree JJ, Roe AJ, Flockhart A, McAteer SP, Xu X, Shaw D, Mahajan A, Beatson SA, Best A, Lotz S, Woodward MJ, La Ragione R, Murphy KC, Leong JM, Gally DL.** 2011. Transcriptional regulators of the GAD acid stress island are carried by effector protein-encoding prophages and indirectly control type III secretion in enterohemorrhagic Escherichia coli O157:H7. Mol. Microbiol. **80:**1349–1365. http://dx.doi.org/10.1111/j.1365-2958.2011.07650.x.
40. **Roe AJ, Naylor SW, Spears KJ, Yull HM, Dransfield TA, Oxford M, McKendrick IJ, Porter M, Woodward MJ, Smith DG, Gally DL.** 2004. Co-ordinate single-cell expression of LEE4- and LEE5-encoded proteins of Escherichia coli O157:H7. Mol. Microbiol. **54:**337–352. http://dx.doi.org/10.1111/j.1365-2958.2004.04277.x.
41. **Besser TE, Shaikh N, Holt NJ, Tarr PI, Konkel ME, Malik-Kale P, Walsh CW, Whittam TS, Bono JL.** 2007. Greater diversity of Shiga toxin-encoding bacteriophage insertion sites among Escherichia coli O157:H7 isolates from cattle than in those from humans. Appl. Environ. Microbiol. **73:**671–679. http://dx.doi.org/10.1128/AEM.01035-06.
42. **Pradel N, Boukhors K, Bertin Y, Forestier C, Martin C, Livrelli V.** 2001. Heterogeneity of Shiga toxin-producing Escherichia coli strains isolated from hemolytic-uremic syndrome patients, cattle, and food samples in central France. Appl. Environ. Microbiol. **67:**2460–2468. http://dx.doi.org/10.1128/AEM.67.6.2460-2468.2001.
43. **Nataro JP, Kaper JB.** 1998. Diarrheagenic Escherichia coli. Clin. Microbiol. Rev. **11:**142–201.
44. **Kenny B, DeVinney R, Stein M, Reinscheid DJ, Frey EA, Finlay BB.** 1997. Enteropathogenic E. coli (EPEC) transfers its receptor for intimate adherence into mammalian cells. Cell **91:**511–520. http://dx.doi.org/10.1016/S0092-8674(00)80437-7.
45. **Jerse AE, Yu J, Tall BD, Kaper JB.** 1990. A genetic locus of enteropathogenic Escherichia coli necessary for the production of attaching and effacing lesions on tissue culture cells. Proc. Natl. Acad. Sci. U. S. A. **87:**7839–7843. http://dx.doi.org/10.1073/pnas.87.20.7839.
46. **Flockhart AF, Tree JJ, Xu X, Karpiyevich M, McAteer SP, Rosenblum R, Shaw DJ, Low CJ, Best A, Gannon V, Laing C, Murphy KC, Leong JM, Schneiders T, La Ragione R, Gally DL.** 2012. Identification of a novel prophage regulator in Escherichia coli controlling the expression of type III secretion. Mol. Microbiol. **83:**208–223. http://dx.doi.org/10.1111/j.1365-2958.2011.07927.x.
47. **Darling AE, Mau B, Perna NT.** 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One **5:**e11147. http://dx.doi.org/10.1371/journal.pone.0011147.

BMC Genomics



Whole genome sequence comparison of *vtx2*-converting phages from Enteroaggregative Haemorrhagic *Escherichia coli* strains

Grande *et al.*

BioMed Central

Grande *et al. BMC Genomics* 2014, **15**:574
http://www.biomedcentral.com/1471-2164/15/574

Grande et al. BMC Genomics 2014, 15:574
http://www.biomedcentral.com/1471-2164/15/574

BMC Genomics

RESEARCH ARTICLE Open Access

Whole genome sequence comparison of *vtx2*-converting phages from Enteroaggregative Haemorrhagic *Escherichia coli* strains

Laura Grande[1,2*], Valeria Michelacci[1], Rosangela Tozzoli[1], Paola Ranieri[1], Antonella Maugliani[1], Alfredo Caprioli[1] and Stefano Morabito[1]

Abstract

Background: Enteroaggregative Haemorrhagic *E. coli* (EAHEC) is a new pathogenic group of *E. coli* characterized by the presence of a *vtx2*-phage integrated in the genomic backbone of Enteroaggregative *E. coli* (EAggEC). So far, four distinct EAHEC serotypes have been described that caused, beside the large outbreak of infection occurred in Germany in 2011, a small outbreak and six sporadic cases of HUS in the time span 1992–2012. In the present work we determined the whole genome sequence of the *vtx2*-phage, termed Phi-191, present in the first described EAHEC O111:H2 isolated in France in 1992 and compared it with those of the *vtx*-phages whose sequences were available.

Results: The whole genome sequence of the Phi-191 phage was identical to that of the *vtx2*-phage P13374 present in the EAHEC O104:H4 strain isolated during the German outbreak 20 years later. Moreover, it was also almost identical to those of the other *vtx2*-phages of EAHEC O104:H4 strains described so far. Conversely, the Phi-191 phage appeared to be different from the *vtx2*-phage carried by the EAHEC O111:H21 isolated in the Northern Ireland in 2012. The comparison of the *vtx2*-phages sequences from EAHEC strains with those from the *vtx*-phages of typical Verocytotoxin-producing *E. coli* strains showed the presence of a 900 bp sequence uniquely associated with EAHEC phages and encoding a tail fiber.

Conclusions: At least two different *vtx2*-phages, both characterized by the presence of a peculiar tail fiber-coding gene, intervened in the emergence of EAHEC. The finding of an identical *vtx2*-phage in two EAggEC strains isolated after 20 years in spite of the high variability described for *vtx*-phages is unexpected and suggests that such *vtx2*-phages are kept under a strong selective pressure.
The observation that different EAHEC infections have been traced back to countries where EAggEC infections are endemic and the treatment of human sewage is often ineffective suggests that such countries may represent the cradle for the emergence of the EAHEC pathotype. In these regions, EAggEC of human origin can extensively contaminate the environment where they can meet free *vtx*-phages likely spread by ruminants excreta.

Keywords: Enteroaggregative haemorrhagic *E. coli*, *vtx*-phages, Whole genome sequence, Tail fibers

* Correspondence: laur.grande@gmail.com
[1]EU Reference Laboratory for E. coli, Veterinary Public Health and Food Safety Department, Istituto Superiore di Sanità, Viale Regina Elena 299, Rome 00161, Italy
[2]Università degli Studi di Roma Tre, Rome, Italy

Background

Diarrheagenic *Escherichia coli* (DEC) are a heterogeneous group of pathogenic *E. coli* causing a wide range of enteric diseases in humans and animals [1].

Enteroaggregative *E. coli* (EAggEC) are a DEC pathotype inducing a gastrointestinal illness characterized by long-lasting watery, mucoid, secretory diarrhoea with low-grade fever and little or no vomiting [2,3]. EAggEC infections are a common cause of acute diarrheal illness among children in low-income countries, but sporadic cases and outbreaks are recorded in industrialized countries as well [4,5].

In 2011, a large *E. coli* outbreak struck Germany causing more than 4,000 human infections including 50 deaths [6]. The outbreak strain, an *E. coli* O104:H4, showed the presence of the typical virulence genes of EAggEC such as *aggR*, *aaiC*, *sepA*, *aatA* and, at the same time, it carried a bacteriophage conveying the genes encoding the Verocytotoxin (Vtx) subtype 2a (*vtx2a*) [7]. In accordance with this genomic asset, the strain showed the Enteroaggregative typical "stacked brick" adhesion to cultured Hep-2 cells and was able to produce Vtx2 [8]. The infecting strain thus displayed an unusual combination of virulence features comprising the colonization repertoire from EAggEC coupled with the production of a toxin typically produced by Vtx-producing *E. coli* (VTEC), a DEC type causing haemorrhagic colitis and Haemolytic Uremic Syndrome (HUS) worldwide [1].

The impact of the German outbreak was so huge that the epidemic strain became iconic of a new DEC type: the Enteroaggregative Haemorrhagic *E. coli* (EAHEC) [9]. The occurrence of the German outbreak also caused the scientific community to look retrospectively at the reported HUS cases linked to infections with atypical VTEC types or to browse the scientific literature in order to assess if other EAHEC cases of infection could be retrieved. It turned out that in the time period 1992–2012 a small outbreak and at least six sporadic cases of HUS had been described as being associated with EAHEC strains belonging to four different serotypes: O111:H2, O86:HNM, O104:H4 and O111:H21 [8,10-13].

The analysis of the whole genome sequence of the EAHEC O104:H4 that caused the German outbreak in 2011 showed that the *vtx2*-phage is inserted in a bacterial genomic backbone typical of EAggEC [14], therefore the EAHEC pathotype seems to have arisen from the acquisition of *vtx2*-phages by classical EAggEC strains.

The appearance of the EAHEC group has shown that the stable acquisition of *vtx*-phages seems to have occurred at least twice by two different DEC groups, the EAggEC and the atypical EPEC (aEPEC) from which the typical VTEC pathotype derives [15-17]. Moreover, the ability of *vtx2*-phages to infect, in the laboratory conditions, different *E. coli* pathogroups including ExPEC has been reported [18,19]. This observation, together with the isolation of Enterobacteriaceae other than *E. coli* producing Vtx from cases of human disease [20,21] suggests that *vtx*-phages can infect a range of bacterial hosts wider than expected, confirming the pivotal role of phages in the evolution of bacterial pathogens.

In the present work we determined the whole genome sequence of the *vtx2*-phage present in the first EAHEC ever described and compared it with that of the *vtx2*-phages present in the EAHEC O104:H4 and O111:H21 available in the public repositories and with those of other *vtx*-phages, with the aim of investigating the mechanisms underlying the evolution of the EAHEC pathotype.

Methods

Bacterial strains

The EAHEC O111:H2 strain ED 191 has been used to obtain the *vtx2*-phage subjected to whole genome sequencing and is part of the collections held at Istituto Superiore di Sanità. The strain's characteristics have been described in a previous publication [10].

E. coli K12 strain LE392 [22] has been used as a propagator strain in infection experiments for the *vtx2*-phage amplification prior to sequencing.

Determination of the vtx2-phages integration sites in the E. coli genome

The *vtx2*-phage integration site in the *E. coli* strain ED 191 has been determined. The occupancy of loci *sbcB, wrbA, yehV*, Z2577, and *yecE* has been assessed as previously described [18].

Infection experiments and phages propagation

The EAHEC strain ED 191 has been exposed to UV light in order to induce the excision of phage genome from the bacterial chromosome [23]. In detail, the bacterial strain has been grown in Luria-Bertrani (LB) broth (Oxoid Limited, Basingstoke Hampshire, UK) overnight at 37°C with vigorous shaking. The culture has been diluted 1:100 in LB modified broth (LB with 0.001% thiamine V/V) and grown to 0.5 OD 600, pelleted and re-suspended in a sterile solution of $CaCl_2$ 10 mM. The culture has been exposed to UV light (130 μJoule X 100) in a crosslinker "Stratalinker® UV crosslinker" (Stratagene Cloning Systems, La Jolla, CA, USA). After induction, the culture has been diluted in LB modified broth and incubated at 37°C for 5 hours with vigorous shaking. The culture has been centrifuged and the supernatant containing phages particles filtered with 0,22 μm pore-filters. 100 μl of phage particles suspension have been added to 100 μl of a culture of the propagator strain *E. coli* LE392 grown in LB modified broth at 0.5 OD 600 and maintained at 37°C for 20 minutes with static incubation. Each tube has been added with 3.5 ml of LB modified soft agar (LB modified broth with agar 7 g/L) at 42°C and immediately

poured on LB modified agar plates (LB modified broth with 15 g/L agar). Plates have been incubated overnight at 37°C.

Four ml of SM buffer (100 mM NaCl, 8 mM $MgSO_4{\cdot}7H_2O$, 50 mM Tris–HCl 1 M pH 7.5, Gelatin 0.002%) have been dispensed to each plate in order to recover phages particles from the lytic plaques and kept overnight at 4°C. The phage suspension in SM has been recovered and chloroform has been added at 5% final concentration. The phage suspension has been centrifuged at 500*xg* 10 minutes twice for removing agar debris and used to re-infect the propagator *E. coli* strain LE392 in the conditions described above in order to increase the phage titre. Finally, the phage suspension has been concentrated by using Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-30 tubes (Merck Millipore, Billerica, MA, USA) with a cut-off of 30 KDa. Final phage titre was 7×10^{10} PFU/ml.

CsCl gradient and viral DNA extraction

The suspension has been purified by Isopycnic Centrifugation through CsCl Equilibrium gradient as described by Sambrook and Russell [23]. Briefly, 2 ml of the phage suspension have been added with 1.5 g of CsCl, transferred to ultracentrifuge tubes, which have been filled with a CsCl solution 0.75 g/ml. The tubes have been finally sealed with mineral oil and centrifuged in a Beckman ultracentrifuge at 154,000*xg*, 8°C for 20 hours in a SW-41 rotor. The band containing the phage particles has been collected with a syringe by puncturing the tube. The recovered solution containing the purified phage particles has been dialyzed in against 10 mM NaCl, 50 mM Tris–HCl pH 8.0, 10 mM $MgCl_2$. A final volume of 1 ml was obtained.

The suspension has been treated by adding 100 units of DNase I RNase-free (New England Biolabs, USA) at 37°C for one hour to eliminate free DNA contaminating the phage suspension. Finally a treatment with proteinase K 50 μg/ml at 56°C for one hour has been carried out to disrupt the phage capsid followed by DNA extraction with phenol-chloroform-isoamyl alcohol [23]. Phage DNA concentration after the purification step was estimated to be 239.4 ng/μl.

Library preparation and whole genome sequencing of the phage DNA

Phage DNA has been sequenced with an Ion Torrent PGM semiconductor sequencer (Life Technologies, Carlsbad, USA) using the 200 bp protocol. An Ion Torrent 314 chip has been used following the manufacturer instructions (Life Technologies, Carlsbad, USA). Genomic library has been obtained by shearing 1 μg of DNA in blunt-ended fragments followed by linking the Ion Adapters using the protocol included in the Ion Xpress™ Plus Fragment Library Kit (Life Technologies, Carlsbad, USA). The sized and ligated fragments have been amplified by emulsion-PCR using the Ion OneTouch 200 Template kit and instruments (Life Technologies, Carlsbad, USA).

Assembly and further bioinformatics analyses

The reads resulting from the sequencing of the *vtx2*-phage DNA from the EAHEC O111:H2 strain, termed Phi-191, have been assembled in contigs by using the open source MIRA software integrated in the Ion Torrent Server. Contigs have been imported in Kodon software (Applied Maths NV, Sint-Martens-Latem, BE) for analysis. To fill in the gaps between contigs, a total of 95 primers have been designed and used for sequencing by Sanger technology using a Genetic analyzer 3130 (Life Technologies, Carlsbad, USA). Mauve software [24] has been used to order the contigs using the sequence of the phage P13374 from the *E. coli* O104:H4 that caused the outbreak in Germany in 2011 [GenBank: NC_018846.1] as reference. The complete sequence of the Phi-191 phage has been annotated by Prokka tool on the online server Galaxy/CRS4 [25] and submitted to GenBank [GenBank: KF971864]. The G + C content has been analysed by the GC calculator free online tool [26]. Identification of putative tRNA genes has been performed using tRNAscan-SE [27].

The raw sequence data (short reads) from the EAHEC O111:H21 strain 226 were retrieved from the SRA database present on NCBI website [NCBI SRA: SRA055981] and aligned on the complete sequence of Phi-191 phage, determined in the present study and used as reference, with the Bowtie2 free software implemented in the Galaxy/CRS4 server [25].

Genomic comparisons between the available *vtx*-phages sequences have been performed by using the BLAST algorithm available at NCBI [28] and the Mauve free alignment software [24]. Comparison map between *vtx*-phages has been generated by Circoletto online tool [29,30]. For the pictogram construction, bit-score values have been used to describe the quality of the alignment at a given point. The bit-score is a normalized version of the score value returned by the BLAST searches, expressed in bits [28].

Ethics

This work does not include animal testing and does not report human data. All the information regarding outbreaks and cases of infection are all from already published papers properly referenced in the text.

Results

Sequencing of the vtx2-phage from the EAHEC O111:H2

The whole genome sequencing of the *vtx2*-phage from the EAHEC O111:H2 strain, Phi-191, produced 320,044 reads of a mean length of 204 bp, for a total of 65.30 Mb sequenced. The assembly of the Phi-191 sequence reads

using as a reference the genome of the P13374 phage [GenBank: NC_018846.1] produced 151 contigs that were further analysed using bioinformatics resources. The estimated coverage of the phage genome was 1088X.

The length of the whole sequence of Phi-191 resulted to be about 61 Kb (61,036 bp) [GenBank: KF971864] with a mean G + C content of 50.2%.

Sequence annotation revealed the presence of 87 predicted coding sequences including the genes encoding the subtype 2a of the Verocytotoxin and three transfer RNAs (tRNAs) (Figure 1).

As it has been reported for the P13374 phage, conveying the *vtx2* genes in the EAHEC O104:H4 strain that caused the German outbreak in 2011 [31], the genome of Phi-191 (Figure 1) only included the lambda genes *cI* and *cro* while the other genes typically composing the regulatory repertoire of lambda phages such as *cII, cIII, N, Ea10* and *gam* seemed not to be present [31]. Most of these genes are involved in the regulation of the switch between lysogeny and lytic cycle in lambda-like phages. Therefore, their absence suggests the existence of an alternative mechanism used to regulate the choice of entry into lysogenic state. The Phi-191 phage is integrated in the bacterial chromosome in the *wrbA* locus, as it has been described for P13374 [31].

Comparison of the Phi-191 whole genome sequence with other vtx-phage sequences

A BLAST search of the Phi-191 *vtx2*-phage whole genome sequence against those collected in the nucleotide repository held at NCBI, returned a 99% identity with that of the P13374 phage from O104:H4 strain CB13374 which caused the German outbreak in 2011 [GenBank: NC_018846.1] and a 91% identity with that of the *vtx2*-phage TL-2011c carried by the VTEC O103:H25 strain that caused a severe HUS outbreak in Norway in 2006 [GenBank:JQ011318] [32]. This finding is in agreement with how reported in the literature for the *vtx*-phage carried by the EAHEC O104:H4 from the German outbreak and the TL-2011c phage [33]. As expected, the Phi-191

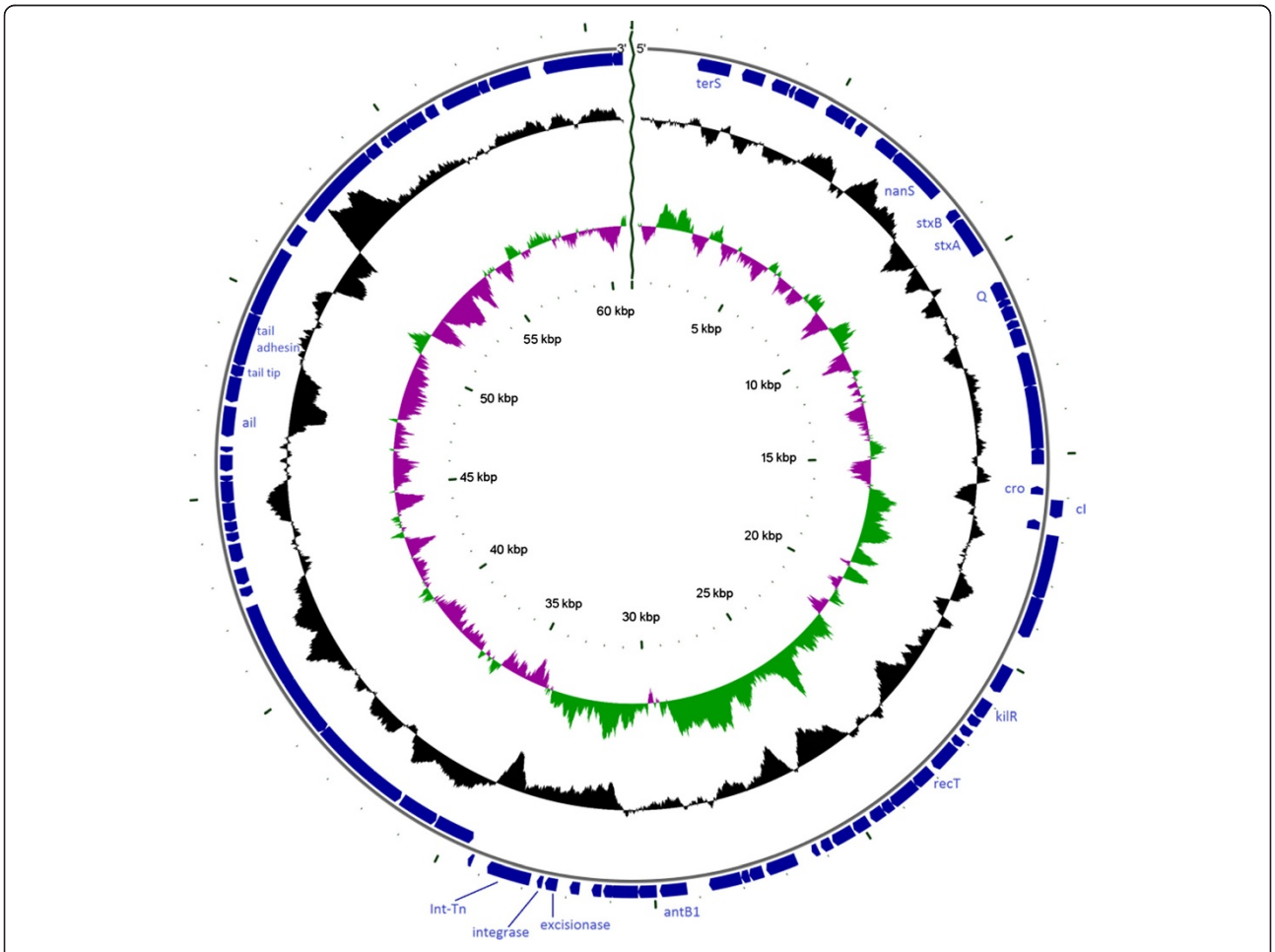


Figure 1 Genomic organization of Phi-191 phage. Coding sequences are represented as blue bars in the outer circle. Putative genes are labelled according to the predicted functions of their products, if known. The whole GC content and GC skew of leading and lagging strand is shown in the black and coloured inner circles, respectively.

phage sequence also showed 99% identity with the *vtx2*-phage from another EAHEC O104:H4 strain isolated during the 2011 German outbreak, the strain 2011C-3493 [GenBank:CP0032891.1] and was highly related (97% nucleotide sequence identity) to the sequence of two *vtx*2-phages from two EAHEC O104:H4 strains isolated from as many haemorrhagic colitis cases occurred in Georgia in 2009 [GenBank:CP003301.1, CP003297.1] [34].

The other scores included hits with query coverage values ranging from 87% down to 60% and having 98%-99% sequence identity, with a number of other *vtx*-phages identified in different VTEC strains (Table 1). The alignment of all the phage sequences comprised in the 100%-60% similarity range is shown in a pictogram generated with the Circoletto online tool [30] (Figure 2) and in another one produced by Mauve software (in Additional file 1: Figure S1).

Recently, another EAHEC strain of serotype O111:H21 (strain 226), isolated from an HUS case occurred in Northern Ireland in 2012, has been described [13]. The short reads from the whole genome-sequencing project of this strain are available in NCBI sequence reads archive [NCBI SRA: SRA055981] and have been used for comparison. Unfortunately, among the 456 contigs obtained from the *de novo* assembly of the reads, the one including the *vtx2* genes was only 8,042 bp long, hindering the analysis of the entire sequence of the *vtx2*-encoding phage. Nevertheless, its whole-genome sequencing reads were aligned to the complete genomic sequence from the Phi-191 phage. The alignment was carried out with Bowtie2 software and visualized on the Integrative Genomics Viewer (IGV) free tool [45] (data not shown). This analysis failed to identify, in the genome of strain 226, a *vtx*2-phage with characteristics similar to that of Phi-191. In particular, the alignment showed that the reads from strain 226 mapped on the sequence of the Phi-191 phage only in the region comprising the Vtx2 A subunit-coding gene and extending 4.2 kb downstream the gene encoding the Vtx2 B subunit. It should be noted that alignments conducted using the whole short reads set could have generated a too high background noise deriving from the presence of sequences from other bacteriophages integrated in different regions of the chromosome of the O111:H21 strain 226. Nevertheless, the lack of sequencing reads aligning to the rest of the sequence of Phi-191 suggests that a different *vtx2* phage intervened in the emergence of this last EAHEC strain.

Table 1 List of the BLAST hits of Phi-191 DNA sequence aligned to the *vtx*-phages sequences from typical VTEC strains

Strain	Similarity%	Acc. No.	Reference
E. coli O145:H28 str. RM13514	87%	CP006027.1	[35]
E. coli O103:H2 str. 12009 DNA	86%	AP010958.1	[35,36]
Phage VT2 phi_272	85%	HQ424691.1	-
E. coli O157:H7 str. TW14359	67%	CP001368.1	[37]
E. coli O157:H7 str. EC4115	67%	CP001164.1	[38]
E. coli O111:H- str. 11128	65%	AP010960.1	[36]
E. coli O157:H7 str. Sakai	65%	BA000007.2	[39]
E. coli Xuzhou21	65%	CP001925.1	[40]
E. coli O157:H7 EDL933	65%	AE005174.2	[41]
Stx2 converting phage II	64%	AP005154.1	[42]
Stx2 converting phage I	63%	AP004402.1	[43]
Enterobacteria phage Min27	63%	EU311208.1	[44]
Stx1 converting phage	60%	AP005153.1	[42]

The similarity score indicates the query coverage values with 98%-99% sequence identity. The hits with similarity values down to 60% are shown.

Genomic comparison of Phi-191 with other vtx2-phages

In order to delve into the genomic organization of the different *vtx*2-phages we performed a progressive Mauve alignment between all the phages sequences used in the nucleotide comparison (Additional file 1: Figure S1). This analysis onfirmed the presence of similarities between the sequences but also highlighted that an extensive rearrangement between the sequence-blocks must have occurred at some points during the phages' evolution. This observation is in agreement with how reported in literature [46-48].

A deeper analysis of the Phi-191 genome was conducted by using phage sequences selected among those from typical VTEC displaying the highest similarity values (Table 1).

The alignment of the Phi-191 sequence with those of phage TL-2011c (O103:H25), and the phages from the VTEC strains RM13514 (O145:H28) and 12009 (O103:H2), the causative agent of a romaine lettuce-associated outbreak occurred in the US in 2010 [35] and isolated in Japan from a sporadic case of diarrhoea [36], respectively, showed the presence of two sequence blocks that seemed to be peculiar to the EAHEC *vtx2*-phage and divergent from the same regions in the three classical VTEC associated-phages (Figure 3). One of the two regions was 1,500 bp long and comprised 140 bp of the 5'-terminus of a gene encoding a lysozyme (*rrrD*), two complete genes encoding the lysis protein S and a hypothetical protein, and 440 bp of the 3'-terminus of a gene annotated as a hypothetical protein. The 1,500 bp region corresponded to the nucleotidic positions 5,240-6,725 in the P13374 genome (ORFs 12–15). The other fragment was 900 bp long and spanned the region comprised between nucleotides 42,160 and 43,050 in the genome of the P13374 phage. The 900 bp fragment shared 100% homology with a region of 730 bp at the 3'-terminus of the ORF65 of P13374 phage encoding a phage tail fiber and a 72 bp fragment of the ORF66 coding for a tail fiber adesine.

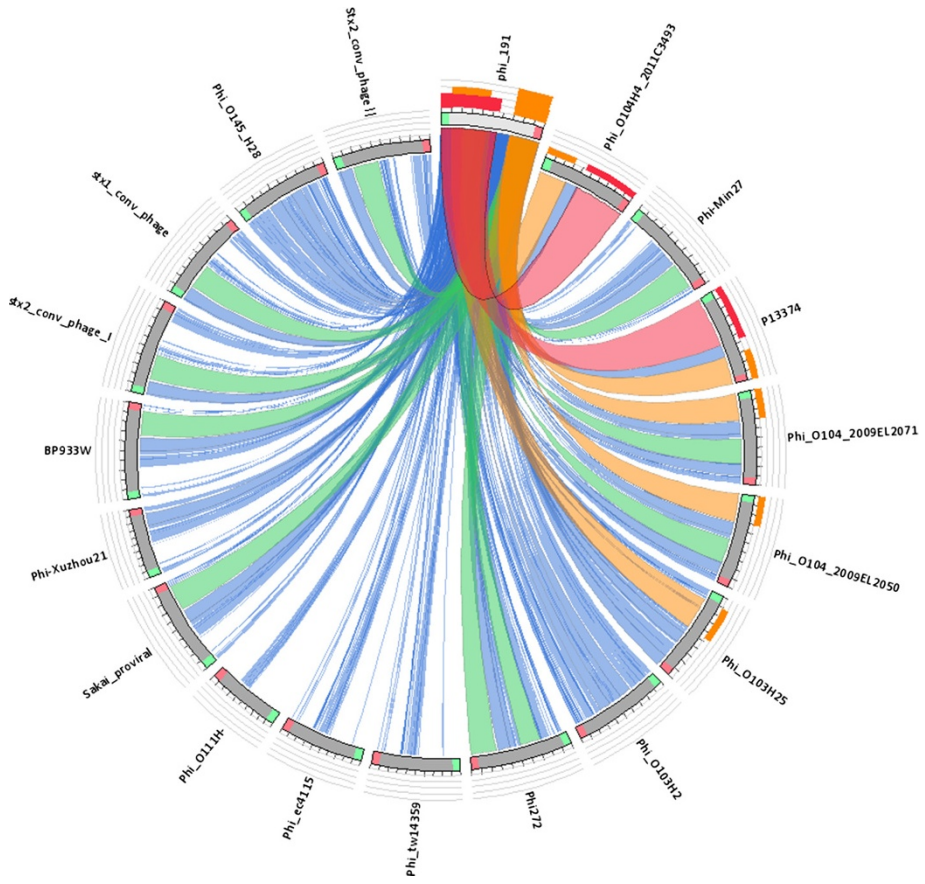


Figure 2 Sequence similarities between Phi-191 and other *vtx*-phages. The picture shows the results of the BLAST local alignments using Phi-191 as a query against the *vtx*-phage sequences with 99% to 60% similarity listed in Table 1. The colours codes blue, green, orange and red represent the overall quality of the aligned segments along the phage sequences, evaluated on the basis of the bit-score values in the worst-to-the-best order (blue to red). The bit-score is a normalized version of the score value returned by the BLAST searches, expressed in bits [28]. The height of the coloured bars in the histogram on the top of the Phi-191 ideogram shows how many times each colour hits a specific fragment of the other phage sequences. A twist in a ribbon indicates that the local alignment is inverted (query and database sequence on opposite strands).

Distribution and analysis of the putative EAHEC-associated sequence regions amongst the vtx-phages

A BLAST comparison was conducted with the aim of investigating the distribution of the two putative EAHEC-associated regions among all the *vtx*-phage sequences available at NCBI. Such an analysis showed that only the 900 bp region encoding phage fiber tail was present in all the fully sequenced *vtx2*-phages from EAHEC strains, [GenBank:CP0032891.1; CP003301.1; CP003297.1] and divergent or absent in all the other phage sequences investigated. Accordingly, the presence of the sole sequence of 900 bp was assessed in the contigs derived from the *de novo* assembly of the short reads from the genome sequence of the O111:H21 strain 226 from Northern Ireland.

Discussion

The *E. coli* genome continuously changes through both small-scale variations and horizontal gene transfer of mobile genetic elements (MGE). Among MGEs, bacteriophages play a pivotal role in the evolution of *E. coli* pathogenic clones [49] by providing a mean for genomic remodelling and conveying important virulence genes such as those encoding the Verocytotoxins (*vtx1* and *vtx2*) in the Verocytotoxin-producing *E. coli* (VTEC).

In 2011 a huge outbreak caused by an Enteroaggregative Haemorrhagic *E. coli* (EAHEC) O104:H4 struck Germany with more than 3,000 cases of infection, 800 HUS, and 50 deaths [12]. The causative agent was a mosaic strain deriving from the lysogenization of an Enteroaggregative *E. coli* strain with a *vtx2*-phage [8]. Such a virulence combination was indeed associated with elevate pathogenicity, as demonstrated by the high rate of human infections evolving to HUS, even among adults (88% and 42 years of median age), and the heavy toll of 50 deaths [6].

This arrangement of virulence factors in *E. coli* strains from human disease had been occasionally reported before during the investigation of a small outbreak of HUS

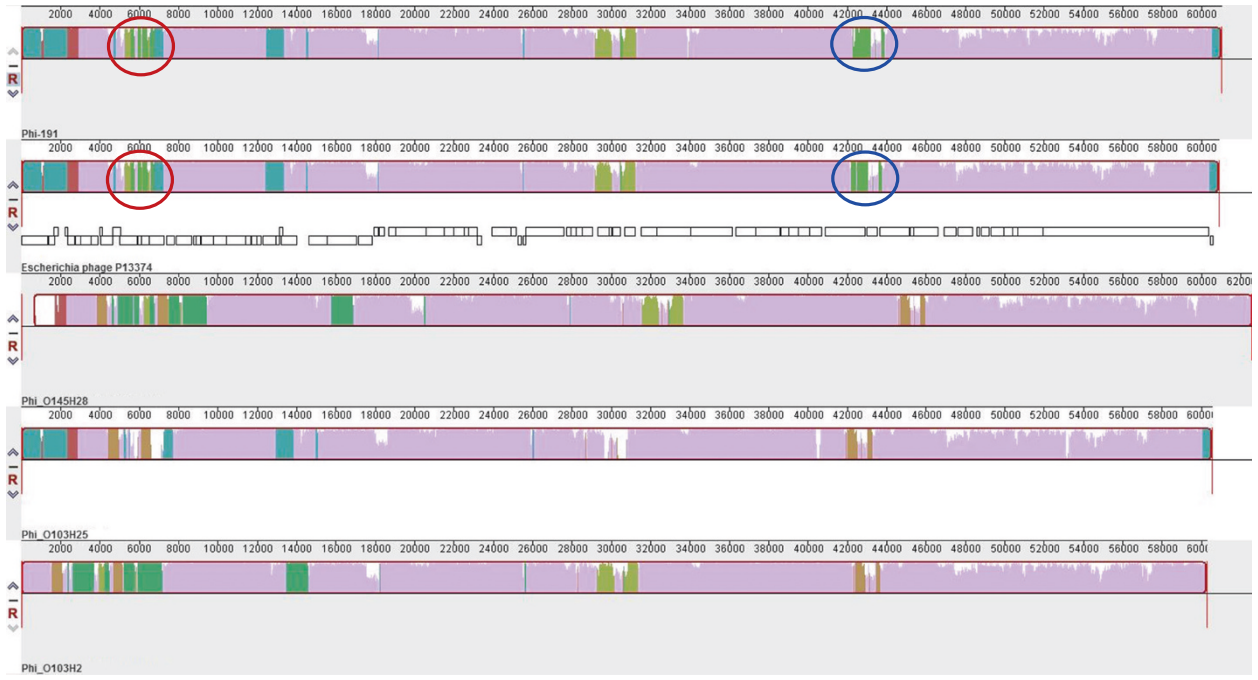


Figure 3 Mauve Progressive Alignment of *vtx2*-phage genomes from EAHEC and VTEC strains showing the EAHEC-specific regions. Blocks with the same colours indicate the *vtx2*-phages regions with identical DNA sequence. White blocks in a phage sequence indicate regions lacking of correspondence in the other sequences. Fragments of the phage genomes that are peculiar to EAHEC *vtx2*-phages are marked with circles: a red circle indicates the sequence encoding a hypothetical protein; the sequence encoding the phage fiber tail is encircled in blue.

occurred in France in 1992 and a case of infection in Japan in 1999 [10,11].

The occurrence of the German outbreak caused the scientific community to look back retrospectively at the repositories of VTEC infections records and culture collections and it turned out that some other sporadic cases of infections with Vtx2-producing Enteroaggregative *E. coli* O104:H4 had already occurred in Europe and Asia in the time span 2001–2011 [12,31]. Finally a HUS case occurred in Northern Ireland in 2012 was demonstrated to be associated with an EAHEC O111:H21 [13].

The observation of sporadic cases and outbreaks occurring throughout a 20-years time span, all caused by Vtx2-producing EAggEC and belonging to different serotypes, strengthens the hypothesis that these pathogenic *E. coli* represent a new pathogroup, as it has been recently proposed [9].

To better understand the events underlying the emergence of EAHEC we determined the whole genome sequence of Phi-191, the *vtx2*-phage present in the EAHEC O111:H2 isolated during the French outbreak of 1992, and compared it with the sequences of the *vtx2*-phage present in the EAHEC strains described in the following years and available in GenBank.

Interestingly, the genomic sequence of Phi-191 was almost identical to that of the *vtx2*-phages from the EAHEC O104:H4 strains isolated during the 2011 German outbreak about 20 years later. This is noteworthy since *vtx*-phages are characterized by a high degree of variability [50,51]. It is conceivable that the same *vtx2*-phage has been acquired in two different events and that the selective pressure impeded the accumulation of changes in the phage sequence before the phage infection events occurred.

However, the EAHEC O111:H21 isolated in Northern Ireland in 2012 seems to host a different type of *vtx2*-phage, suggesting that at least two different *vtx2*-phage types have been successfully transferred to EAggEC. Unfortunately, the sequence of the phage of the EAHEC O86: HNM isolated in Japan in 1999 was not available for comparison [11].

It has been hypothesized that the infection with a lambdoid phage can be mediated by the cross-talking between the bacterium and the phage resulting in host specificity [52]. An extended comparison of the EAHEC *vtx*2-phages with the whole genome sequences of *vtx*-phages from VTEC strains available at NCBI returned a wide range of similarities between sequences, going from 87% to 60% and lower. This picture is in line with how reported for the general variability of *vtx*-phages sequences [50]. Interestingly, one region of 900 bp, identified in the Phi-191 and encoding a tail fiber, was present in all the *vtx*2-phages from EAHEC and was also present in the short reads dataset from the EAHEC O111:H21. At the same time this DNA sequence was

absent in all the *vtx*-phage sequences identified in VTEC strains and stored at NCBI.

This is in agreement with previously reported data which pointed at a larger fragment including this region as one of those differentiating the P13374 genome from the *E. coli* phage TL-2011c (O103:H25) and not exhibiting significant homology to known *vtx*-encoding phages [31].

The annotation of the Phi-191 genome showed that this sequence peculiar to EAHEC *vtx2*-phages contains part of a gene encoding a type of phage tail fiber displaying some conserved aminoacidic motifs such as a Collagen triple helix repeat (20 copies) [NCBI CDD:189968] and the Peptidase_S74 [NCBI CDD:258151]. The latter is the C-terminal domain of the bacteriophage protein endosialidase, which forms homotrimeric molecules and releases itself from the end-tail-spike of the bacteriophages [53].

The 900 bp-long sequence could potentially encode part of the mechanism defining the specificity of the *vtx*2-phages for EAggEC strains, being directly involved in the phage-bacterium interaction. As a matter of fact, several authors reported that the interactions between phage tail fibers and host proteins, such as LamB and OmpC [52,53] contribute to the success of the infection, as demonstrated by the finding that *lamB* gene mutations block phage adsorption [52]. It is therefore conceivable that differences in phage tail fibers may contribute to define *vtx2*-phages tropism for *E. coli* recipients, although this hypothesis together with the mechanisms underlying this process still need to be verified.

For a successful infection to occur, suitable *vtx*-phages and *E. coli* acceptors need to meet in the same environment. In the case of the emergence of typical VTEC, the events of *vtx*-phage acquisition probably occurred at the level of the gastrointestinal tract of ruminants [54] where both *vtx*-phages and aEPEC are abundant [55,56].

Conversely, the EAHEC emergence is probably not directly connected to an animal reservoir since EAggEC are human pathogens with an inter-human transmission of the infection [1]. The environment, in turn, plays a role in the pathogen's amplification cycles, particularly in geographic areas characterised by poor hygienic conditions, where the lack of effective human sewage treatments make the infections with enteric pathogens, including EAggEC, endemic [57]. In such a scenario, an environment contaminated with ruminant's excreta might have been the source of the *vtx2*-phages found in EAHEC as it has been recently proposed [16]. Such a picture may account for the existence of a favourable setting for the EAggEC and the *vtx*-phages to come in contact and for the following selection process resulting in the occasional emergence of an *E. coli* strain matching the EAHEC definition.

Conclusions

The *vtx2*-phages characterising EAHEC seem to belong to a sub-population of *vtx*-phages kept under selective pressure and characterised by the presence of a gene encoding a tail fiber, which could be involved in the mechanism used to recognize the EAggEC. The new EAHEC pathogroup may have emerged following multiple *vtx2*-phage acquisition events favoured by an overlapping of a human reservoir of pathogenic *E. coli*, the EAggEC, with the known animal reservoir of *vtx*-phages. The emergence of this new *E. coli* pathogroup further witnesses the great adaptability and plasticity of this bacterial species and underlines the need to rethink the global asset of hygienic practices to mitigate enteric infections worldwide; particularly in the presence of a global market of foodstuffs that is extending its boundaries towards low-income countries in the quest of new sources to meet the always increasing demand of cheap and exotic food commodities.

Additional file

Additional file 1: Figure S1. Mauve Progressive Alignment of Phi-191 genome with *vtx2*-phage genomes from EAHEC and VTEC strains showing the highest score of similarity. Blocks with the same colours indicate the *vtx2*-phages regions with identical DNA sequence. White fragments in a phage sequence indicate regions lacking of correspondence in the other sequences. Connecting lines link the same genomic block in different genomes and help to pinpoint the re-arrangement between *vtx*-phage genomes.

Abbreviations

aEPEC: Atypical enteropathogenic *Escherichia coli*; DEC: Diarrheagenic *Escherichia coli*; EAggEC: Enteroaggregative *Escherichia coli*; EAHEC: Enteroaggregative haemorrhagic *Escherichia coli*; HC: Haemorrhagic colitis; HUS: Haemolytic uremic syndrome; MGE: Mobile genetic elements; VTEC: Verocytotoxin-producing *Escherichia coli*; vtx: Verocytotoxin genes.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

GL carried out the infection experiments, contributed to the sequencing and drafted the manuscript; MV performed the whole genome sequencing and bioinformatics analysis; TR contributed to all the steps of the whole genome sequencing and revised critically the manuscript; RP carried out part of infection experiments and CsCL gradient; MA helped in the revision of the draft manuscript; CA contributed to the revision of the draft manuscript; MS conceived the study and strongly contributed to revise the manuscript. Finally, all the authors approved the manuscript to be published.

Acknowledgements

We thank Prof. Muniesa for hosting the corresponding author at University of Barcelona, Spain, where the basic techniques for working with phages have been acquired.

Received: 15 April 2014 Accepted: 3 July 2014
Published: 8 July 2014

References

1. Nataro JP, Kaper JB: **Diarrheagenic Escherichia coli.** *Clin Microbiol Rev* 1998, **11**(1):142–201.
2. Bhan MK, Khoshoo V, Sommerfelt H, Raj P, Sazawal S, Srivastava R: **Enteroaggregative Escherichia coli and Salmonella associated with nondysenteric persistent diarrhea.** *Pediatr Infect Dis J* 1989, **8**(8):499–502.

3. Paul M, Tsukamoto T, Ghosh AR, Bhattacharya SK, Manna B, Chakrabarti S, Nair GB, Sack DA, Sen D, Takeda Y: **The significance of enteroaggregative Escherichia coli in the etiology of hospitalized diarrhoea in Calcutta, India and the demonstration of a new honey-combed pattern of aggregative adherence.** *FEMS Microbiol Lett* 1994, **117**(3):319–325.
4. Huang DB, Nataro JP, DuPont HL, Kamat PP, Mhatre AD, Okhuysen PC, Chiang T: **Enteroaggregative Escherichia coli is a cause of acute diarrheal illness: a meta-analysis.** *Clin Infect Dis* 2006, **43**(5):556–563.
5. Estrada-Garcia T, Navarro-Garcia F: **Enteroaggregative Escherichia coli pathotype: a genetically heterogeneous emerging foodborne enteropathogen.** *FEMS Immunol Med Microbiol* 2012, **66**(3):281–298.
6. Frank C, Werber D, Cramer JP, Askar M, Faber M, Heiden M, Bernard H, Fruth A, Prager R, Spode A, Wadl M, Zoufaly A, Jordan S, Stark K, Krause G: **Epidemic profile of Shiga-toxin-producing Escherichia coli O104:H4 outbreak in Germany.** *N Engl J Med* 2011, **365**(19):1771–1780.
7. Bielaszewska M, Mellmann A, Zhang W, Köck R, Fruth A, Bauwens A, Peters G, Karch H: **Characterisation of the Escherichia coli strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study.** *Lancet Infect Dis* 2011, **11**(9):671–676.
8. Scheutz F, Nielsen EM, Frimodt-Møller J, Boisen N, Morabito S, Tozzoli R, Nataro JP, Caprioli A: **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* 2011, **16**(24). Available online: http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19889.
9. Brzuszkiewicz E, Thürmer A, Schuldes J, Leimbach A, Liesegang H, Meyer FD, Boelter J, Petersen H, Gottschalk G, Daniel R: **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* 2011, **193**(12):883–891.
10. Morabito S, Karch H, Mariani-Kurkdjian P, Schmidt H, Minelli F, Bingen E, Caprioli A: **Enteroaggregative, Shiga toxin-producing Escherichia coli O111:H2 associated with an outbreak of hemolytic-uremic syndrome.** *J Clin Microbiol* 1998, **36**(3):840–842.
11. Iyoda S, Tamura K, Itoh K, Izumiya H, Ueno N, Nagata K, Togo M, Terajima J, Watanabe H: **Inducible stx2 phages are lysogenized in the enteroaggregative and other phenotypic Escherichia coli O86: HNM isolated from patients.** *FEMS Microbiol Lett* 2000, **191**(1):7–10.
12. Scavia G, Morabito S, Tozzoli R, Michelacci V, Marziano ML, Minelli F, Ferreri C, Paglialonga F, Edefonti A, Caprioli A: **Similarity of Shiga toxin-producing Escherichia coli O104:H4 strains from Italy and Germany.** *Emerg Infect Dis* 2011, **17**(10):1957–1958.
13. Dallman T, Smith GP, O'Brien B, Chattaway MA, Finlay D, Grant KA, Jenkins C: **Characterization of a verocytotoxin-producing enteroaggregative Escherichia coli serogroup O111:H21 strain associated with a household outbreak in Northern Ireland.** *J Clin Microbiol* 2012, **50**(12):4116–4119.
14. Mellmann A, Harmsen D, Cummings CA, Zentz EB, Leopold SR, Rico A, Prior K, Szczepanowski R, Ji Y, Zhang W, McLaughlin SF, Henkhaus JK, Leopold B, Bielaszewska M, Prager R, Brzoska PM, Moore RL, Guenther S, Rothberg JM, Karch H: **Prospective genomic characterization of the German enterohemorrhagic Escherichia coli O104:H4 outbreak by rapid next generation sequencing technology.** *PLoS One* 2011, **6**(7):e22751.
15. Bielaszewska M, Prager R, Köck R, Mellmann A, Zhang W, Tschäpe H, Tarr PI, Karch H: **Shiga toxin gene loss and transfer in vitro and in vivo during enterohemorrhagic Escherichia coli O26 infection in humans.** *Appl Environ Microbiol* 2007, **73**(10):3144–3150.
16. Beutin L, Hammerl JA, Reetz J, Strauch E: **Shiga toxin-producing Escherichia coli strains from cattle as a source of the Stx2a bacteriophages present in enteroaggregative Escherichia coli O104:H4 strains.** *Int J Med Microbiol* 2013, **303**(8):595–602.
17. Wick LM, Qi W, Lacher DW, Whittam TS: **Evolution of genomic content in the stepwise emergence of Escherichia coli O157:H7.** *J Bacteriol* 2005, **187**(5):1783–1791.
18. Tozzoli R, Grande L, Michelacci V, Ranieri P, Maugliani A, Caprioli A, Stefano M: **Shiga toxin-converting phages and the emergence of new pathogenic Escherichia coli: a world in motion.** *Front Cell Infect Microbiol* 2014, **4**:80. doi:10.3389/fcimb.2014.00080.
19. Schmidt H, Bielaszewska M, Karch H: **Transduction of enteric Escherichia coli isolates with a derivative of Shiga toxin 2-encoding bacteriophage phi3538 isolated from Escherichia coli O157:H7.** *Appl Environ Microbiol* 1999, **65**(9):3855–3861.
20. Tschape H, Prager R, Streckel W, Fruth A, Tietze E, Böhme G: **Verotoxinogenic Citrobacter freundii associated with severe gastroenteritis and cases of haemolytic uraemic syndrome in a nursery school: green butter as the infection source.** *Epidemiol Infect* 1995, **114**(3):441–450.
21. Paton AW, Paton JC: **Enterobacter cloacae producing a Shiga-like toxin II-related cytotoxin associated with a case of hemolytic-uremic syndrome.** *J Clin Microbiol* 1996, **34**(2):463–465.
22. Roberts I, Mountford R, High N, Bitter-Suermann D, Jann K, Timmis K, Boulnois G: **Molecular cloning and analysis of genes for production of K5, K7, K12, and K92 capsular polysaccharides in Escherichia coli.** *J Bacteriol* 1986, **168**(3):1228–1233.
23. Sambrook JRD (Ed): *Molecular Cloning: A laboratory manual.* 3rd edition. New York: Cold Spring Harbor Laboratory Press; 2001.
24. Darling AE, Mau B, Perna NT: **ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement.** *PLoS One* 2010, **5**(6):e11147.
25. **Galaxy/CRS4.** In [http://orione.crs4.it/]
26. **GC Calculator.** In [www.genomicsplace.com/gc_calc.html]
27. Lowe TM, Eddy SR: **tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence.** *Nucleic Acids Res* 1997, **25**(5):955–964.
28. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**(3):403–410.
29. Darzentas N: **Circoletto: visualizing sequence similarity with Circos.** *Bioinformatics* 2010, **26**(20):2620–2621.
30. **Circoletto.** In [http://tools.bat.infspire.org/circoletto/]
31. Beutin L, Hammerl JA, Strauch E, Reetz J, Dieckmann R, Kelner-Burgos Y, Martin A, Miko A, Strockbine NA, Lindstedt BA, Horn D, Monse H, Huettel B, Müller I, Stüber K, Reinhardt R: **Spread of a distinct Stx2-encoding phage prototype among Escherichia coli O104:H4 strains from outbreaks in Germany, Norway, and Georgia.** *J Virol* 2012, **86**(19):10444–10455.
32. Schimmer B, Nygard K, Eriksen HM, Lassen J, Lindstedt BA, Brandal LT, Kapperud G, Aavitsland P: **Outbreak of haemolytic uraemic syndrome in Norway caused by stx2-positive Escherichia coli O103:H25 traced to cured mutton sausages.** *BMC Infect Dis* 2008, **8**:41.
33. L'Abée-Lund TM, Jørgensen HJ, O'Sullivan K, Bohlin J, Ligård G, Granum PE, Lindbäck T: **The highly virulent 2006 Norwegian EHEC O103:H25 outbreak strain is related to the 2011 German O104:H4 outbreak strain.** *PLoS One* 2012, **7**(3):e31413.
34. Ahmed SA, Awosika J, Baldwin C, Bishop-Lilly KA, Biswas B, Broomall S, Chain PS, Chertkov O, Chokoshvili O, Coyne S, Davenport K, Detter JC, Dorman W, Erkkila TH, Folster JP, Frey KG, George M, Gleasner C, Henry M, Hill KK, Hubbard K, Insalaco J, Johnson S, Kitzmiller A, Krepps M, Lo CC, Luu T, McNew LA, Minogue T, Munk CA, *et al*: **Genomic comparison of Escherichia coli O104:H4 isolates from 2009 and 2011 reveals plasmid, and prophage heterogeneity, including shiga toxin encoding phage stx2.** *PLoS One* 2012, **7**(11):e48228.
35. Cooper KK, Mandrell RE, Louie JW, Korlach J, Clark TA, Parker CT, Huynh S, Chain PS, Ahmed S, Carter MQ: **Comparative genomics of enterohemorrhagic Escherichia coli O145:H28 demonstrates a common evolutionary lineage with Escherichia coli O157:H7.** *BMC Genomics* 2014, **15**:17.
36. Ogura Y, Ooka T, Iguchi A, Toh H, Asadulghani M, Oshima K, Kodama T, Abe H, Nakayama K, Kurokawa K, Tobe T, Hattori M, Hayashi T: **Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic Escherichia coli.** *Proc Natl Acad Sci U S A* 2009, **106**(42):17939–17944.
37. Kulasekara BR, Jacobs M, Zhou Y, Wu Z, Sims E, Saenphimmachak C, Rohmer L, Ritchie JM, Radey M, McKevitt M, Freeman TL, Hayden H, Haugen E, Gillett W, Fong C, Chang J, Beskhlebnaya V, Waldor MK, Samadpour M, Whittam TS, Kaul R, Brittnacher M, Miller SI: **Analysis of the genome of the Escherichia coli O157:H7 2006 spinach-associated outbreak isolate indicates candidate genes that may enhance virulence.** *Infect Immun* 2009, **77**(9):3713–3721.
38. Eppinger M, Mammel MK, Leclerc JE, Ravel J, Cebula TA: **Genomic anatomy of Escherichia coli O157:H7 outbreaks.** *Proc Natl Acad Sci U S A* 2011, **108**(50):20142–20147.
39. Makino K, Yokoyama K, Kubota Y, Yutsudo CH, Kimura S, Kurokawa K, Ishii K, Hattori M, Tatsuno I, Abe H, Iida T, Yamamoto K, Onishi M, Hayashi T, Yasunaga T, Honda T, Sasakawa C, Shinagawa H: **Complete nucleotide**

sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7 derived from the Sakai outbreak. *Genes Genet Syst* 1999, **74**(5):227–239.
40. Xiong Y, Wang P, Lan R, Ye C, Wang H, Ren J, Jing H, Wang Y, Zhou Z, Bai X, Cui Z, Luo X, Zhao A, Wang Y, Zhang S, Sun H, Wang L, Xu J: **A novel Escherichia coli O157:H7 clone causing a major hemolytic uremic syndrome outbreak in China.** *PLoS One* 2012, **7**(4):e36144.
41. Perna NT, 3rd Plunkett G, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, Pósfai G, Hackett J, Klink S, Boutin A, Shao Y, Miller L, Grotbeck EJ, Davis NW, Lim A, Dimalanta ET, Potamousis KD, Apodaca J, Anantharaman TS, Lin J, Yen G, Schwartz DC, Welch RA, Blattner FR: **Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.** *Nature* 2001, **409**(6819):529–533.
42. Sato T, Shimizu T, Watarai M, Kobayashi M, Kano S, Hamabata T, Takeda Y, Yamasaki S: **Genome analysis of a novel Shiga toxin 1 (Stx1)-converting phage which is closely related to Stx2-converting phages but not to other Stx1-converting phages.** *J Bacteriol* 2003, **185**(13):3966–3971.
43. Sato T, Shimizu T, Watarai M, Kobayashi M, Kano S, Hamabata T, Takeda Y, Yamasaki S: **Distinctiveness of the genomic sequence of Shiga toxin 2-converting phage isolated from Escherichia coli O157:H7 Okayama strain as compared to other Shiga toxin 2-converting phages.** *Gene* 2003, **309**(1):35–48.
44. Su LK, Lu CP, Wang Y, Cao DM, Sun JH, Yan YX: **[Lysogenic infection of a Shiga toxin 2-converting bacteriophage changes host gene expression, enhances host acid resistance and motility].** *Mol Biol (Mosk)* 2010, **44**(1):60–73.
45. Thorvaldsdóttir H, Robinson JT, Mesirov JP: **Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.** *Brief Bioinform* 2013, **14**(2):178–192.
46. Smith DL, Rooks DJ, Fogg PC, Darby AC, Thomson NR, McCarthy AJ, Allison HE: **Comparative genomics of Shiga toxin encoding bacteriophages.** *BMC Genomics* 2012, **13**:311.
47. Johansen BK, Wasteson Y, Granum PE, Brynestad S: **Mosaic structure of Shiga-toxin-2-encoding phages isolated from Escherichia coli O157:H7 indicates frequent gene exchange between lambdoid phage genomes.** *Microbiology* 2001, **147**(Pt 7):1929–1936.
48. Steyert SR, Sahl JW, Fraser CM, Teel LD, Scheutz F, Rasko DA: **Comparative genomics and stx phage characterization of LEE-negative Shiga toxin-producing Escherichia coli.** *Front Cell Infect Microbiol* 2012, **2**:133.
49. Allison HE: **Stx-phages: drivers and mediators of the evolution of STEC and STEC-like pathogens.** *Future Microbiol* 2007, **2**(2):165–174.
50. Muniesa M, Blanco JE, De Simón M, Serra-Moreno R, Blanch AR, Jofre J: **Diversity of stx2 converting bacteriophages induced from Shiga-toxin-producing Escherichia coli strains isolated from cattle.** *Microbiology* 2004, **150**(Pt 9):2959–2971.
51. García-Aljaro C, Muniesa M, Jofre J, Blanch AR: **Genotypic and phenotypic diversity among induced, stx2-carrying bacteriophages from environmental Escherichia coli strains.** *Appl Environ Microbiol* 2009, **75**(2):329–336.
52. Werts C, Michel V, Hofnung M, Charbit A: **Adsorption of bacteriophage lambda on the LamB protein of Escherichia coli K-12: point mutations in gene J of lambda responsible for extended host range.** *J Bacteriol* 1994, **176**(4):941–947.
53. Blasche S, Wuchty S, Rajagopala SV, Uetz P: **The protein interaction network of bacteriophage lambda with its host, Escherichia coli.** *J Virol* 2013, **87**(23):12745–12755.
54. Kyle JL, Cummings CA, Parker CT, Quiñones B, Vatta P, Newton E, Huynh S, Swimley M, Degoricija L, Barker M, Fontanoz S, Nguyen K, Patel R, Fang R, Tebbs R, Petrauskene O, Furtado M, Mandrell RE: **Escherichia coli serotype O55:H7 diversity supports parallel acquisition of bacteriophage at Shiga toxin phage insertion sites during evolution of the O157:H7 lineage.** *J Bacteriol* 2012, **194**(8):1885–1896.
55. Imamovic L, Ballesté E, Jofre J, Muniesa M: **Quantification of Shiga toxin-converting bacteriophages in wastewater and in fecal samples by real-time quantitative PCR.** *Appl Environ Microbiol* 2010, **76**(17):5693–5701.
56. Hornitzky MA, Mercieca K, Bettelheim KA, Djordjevic SP: **Bovine feces from animals with gastrointestinal infections are a source of serologically diverse atypical enteropathogenic Escherichia coli and Shiga toxin-producing E. coli strains that commonly possess intimin.** *Appl. Environ. Microbiol* 2005, **71**:3405–3412. doi:10.1128/AEM.71.7.3405-3412.2005.
57. Boisen N, Nataro JP and Krogfelt KA: **Enteroaggregative Escherichia coli and Disease**. In Pathogenic *Escherichia coli*, Molecular and Cellular Microbiology. Edited by Morabito S. Caister Academic Press; 2014:217–230.

doi:10.1186/1471-2164-15-574
Cite this article as: Grande *et al.*: Whole genome sequence comparison of *vtx2*-converting phages from Enteroaggregative Haemorrhagic *Escherichia coli* strains. *BMC Genomics* 2014 **15**:574.

CHAPTER 6

Study of the VT-phages host range and the emergence of new pathogenic *Escherichia coli* clones with augmented virulence

Introduction

VTEC pathogenicity relies upon the production of Verocytotoxins (VTs) and the ability to colonize the intestinal mucosa of the host. Some VTEC strains produce a typical histopathological lesion to the enterocyte called Attaching and Effacing (A/E) and have been historically associated with the most severe forms of infections, such as the haemorrhagic colitis (HC) and the life-threatening haemolytic-uremic syndrome (HUS). This paradigm has been subverted by the report of the large and severe HUS outbreak occurred in Germany and France in 2011 caused by a VTEC displaying the colonization machinery of Enteroaggregative *E. coli*. The infectious agent indeed was able to produce VT but presented the typical "stacked-brick adhesion" of EAggEC strains instead of the A/E lesion. One of the proposed nomenclatures for strains presenting this peculiar virulence asset is Enteroaggregative Haemorrhagic *E. coli* (EAHEC). It has to be noted that *E. coli* strains matching this virulence profile had been described before 2011 as the cause of a small HUS outbreak occurred in France in the early '90s. Additionally, few sporadic cases of infection caused by similar EAHEC strains have been retrospectively described in the period 2000-2010. Finally, two more EAHEC strains have been isolated in 2012 and 2013 from a sporadic case of infection in Northern Ireland and a small HUS outbreak in Italy, respectively. These observations, together with the reported isolation of VT-producing Extraintestinal Pathogenic *E. coli* and other Enterobacteriaceae that were able to produce the toxin from severe bacteraemia and HUS, respectively, suggests that VT-phages can infect a range of bacterial hosts wider than expected. The susceptibility of *E. coli* to infection with VT-phages has been experimentally evaluated by different authors with contradictory results and in most cases involved the use of one single VT-phage. In the present work we investigated the ability of a panel of six VT2-phages to infect and lysogenize *E. coli* strains belonging to all the known pathogroups, including diarrheagenic (DEC) and Extraintestinal pathogenic *E. coli* (ExPEC), as well as *E. coli* strains isolated from healthy individuals.

Results

Three of the VT-phages used in this study were isolated from three different EAHEC strains, one was obtained from a hybrid ETEC/VTEC strain and two from typical VTEC strains, an O157 that caused an HUS case in Italy and an O103:H25 responsible of an outbreak in Norway. The panel of *E. coli* used for the susceptibility studies was made up of 35 strains, including tEPEC and aEPEC, ETEC, EIEC, EAggEC, ExPEC, and non-pathogenic *E. coli*. The VT2-phage from the ETEC/VTEC presented the widest host range infecting all the strains in the panel, while the other VT2-phages induced visible lysis on Diarrheagenic and non-pathogenic *E. coli,* only. The ability of the infected strains to produce VT was also investigated by Vero Cell Assay. We found that most of the strains were able to induce a specific cytopathic effect after the infection, also in the absence of visible lytic plaques, suggesting that in these strains the VT-phage was, at least temporarily, integrated in the host chromosome.

In one case only it was possible to isolate an *E. coli* recipient strain stably lysogenized with the VT-phage after several cultural passages, suggesting that a stable acquisition of a VT-phage by *E. coli* although rarely, does occur.

Conclusions

Most of the published experimental findings about the *E. coli* susceptibility to VT-phages suggest that any pathogenic *E. coli* might experience an augmented pathogenicity variant following the stable acquisition of a VT-phage threatening the public health system. Such a hypothesis is supported by the observation of *E. coli* other than typical VTEC and other Enterobacteriaceae that have naturally acquired and stabilized VT-phages into their own chromosome and have been isolated from severe cases of human disease.

Interestingly, all the *E. coli* pathotypes, with the exception of typical VTEC, are characterised by an inter-human transmission of the infections, while the latter and the related VT-phages are reported to have their reservoir in the gastro-intestinal tract of ruminants and thus displaying a zoonotic cycle of the transmission

of the infections. This implies that for an event of VT-phage acquisition to occur, phages and recipient strains should come into contact and interact in a common environment. This niche could be represented by the intestine of mammalian host or, more probably, the environment, stressing the need for taking measures not to favour such an overlap in order to mitigate the risk of the emergence of new pathogenic *E. coli* strains with shuffled virulence features.

Candidate's contributions to the present work:
The candidate's contributions are described in a specific section in the respective paper.

frontiers in
CELLULAR AND INFECTION MICROBIOLOGY

ORIGINAL RESEARCH ARTICLE
published: 20 June 2014
doi: 10.3389/fcimb.2014.00080

Shiga toxin-converting phages and the emergence of new pathogenic *Escherichia coli*: a world in motion

Rosangela Tozzoli, Laura Grande, Valeria Michelacci, Paola Ranieri, Antonella Maugliani, Alfredo Caprioli and Stefano Morabito*

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

Edited by:
Eelco Franz, Centre for Infectious Disease Control, Netherlands

Reviewed by:
Maite Muniesa, University of Barcelona, Spain
Lothar Beutin, Federal Institute for Risk Assessment, Germany

***Correspondence:**
Stefano Morabito, European Reference Laboratory for Escherichia coli, Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare, Istituto Superiore di Sanità, Viale Regina Elena 299, Rome 00161, Italy
e-mail: stefano.morabito@iss.it

Shiga toxin (Stx)-producing *Escherichia coli* (STEC) are pathogenic *E. coli* causing diarrhea, hemorrhagic colitis (HC) and hemolytic uremic syndrome (HUS). STEC are characterized by a constellation of virulence factors additional to Stx and have long been regarded as capable to cause HC and HUS when possessing the ability of inducing the attaching and effacing (A/E) lesion to the enterocyte, although strains isolated from such severe infections sometimes lack this virulence feature. Interestingly, the capability to cause the A/E lesion is shared with another *E. coli* pathogroup, the Enteropathogenic *E. coli* (EPEC). In the very recent times, a different type of STEC broke the scene causing a shift in the paradigm for HUS-associated STEC. In 2011, a STEC O104:H4 caused a large outbreak with more than 800 HUS and 50 deaths. Such a strain presented the adhesion determinants of Enteroaggregative *E. coli* (EAggEC). We investigated the possibility that, besides STEC and EAggEC, other pathogenic *E. coli* could be susceptible to infection with s*tx*-phages. A panel of *stx2*-phages obtained from STEC isolated from human disease was used to infect experimentally *E. coli* strains representing all the known pathogenic types, including both diarrheagenic *E. coli* (DEC) and extra-intestinal pathogenic *E. coli* (ExPEC). We observed that all the *E. coli* pathogroups used in the infection experiments were susceptible to the infection. Our results suggest that the *stx2*-phages used may not have specificity for *E. coli* adapted to the intestinal environment, at least in the conditions used. Additionally, we could only observe transient lysogens suggesting that the event of stable *stx2*-phage acquisition occurs rarely.

Keywords: *Escherichia coli*, Shigatoxin, *stx*-phages, STEC, pathogroups

INTRODUCTION

The ability to produce Shiga toxins (Stx) is the major virulence feature of Shiga toxin-producing *Escherichia coli* (STEC). These potent cytotoxins block the protein synthesis by inactivating ribosomes (Okuda et al., 2006) and their action on the target cells is responsible for the most severe forms of STEC-induced disease, such as hemorrhagic colitis and the life-threatening hemolytic uremic syndrome (HUS) (Karmali, 2009). The Stx-coding genes, *stx*, are conveyed by lambdoid bacteriophages (O'brien et al., 1984), which, following infection of a susceptible *E. coli* strain, are maintained into a lysogenic state in the host chromosome, becoming a virulence marker for STEC. In spite of the striking effect exerted by Stx on susceptible cells, their sole production seems to be not sufficient for STEC to induce severe disease in humans. In fact, STEC associated with HUS usually colonize the gastrointestinal tract of the host by inducing a typical lesion to the enterocyte known as attaching and effacing (A/E) (Mcdaniel and Kaper, 1997). Noteworthy, the ability to induce the A/E lesion is shared with Enteropathogenic *E. coli* (EPEC), another *E. coli* pathogroup historically causing outbreaks of infection with high mortality rates in Europe and US up to the end of the second World War, and nowadays representing a leading cause of diarrhea and infant mortality in the developing countries (Tozzoli and Scheutz, 2014). Until recently this combination of virulence traits, the ability to produce Stx and to cause the A/E lesion, represented the paradigm for HUS-associated STEC, altogether defined as Enterohaemorrhagic *E. coli* (EHEC) (Levine, 1987).

In 2011, in Germany, an outbreak of STEC infections occurred that caused more than 4000 infections, including 900 HUS cases and 50 deaths (Bielaszewska et al., 2011). The outbreak strain, beside the production of Stx, possessed the intestinal colonization apparatus typical of another *E. coli* pathogroup, known as Enteroaggregative *E. coli* (EAggEC). As a matter of fact, the outbreak strain possessed the adhesion-associated genes typical of EAggEC such as *aggR*, *aaiC*, *sepA*, *aatA* (Schmidt et al., 1995; Boisen et al., 2008, 2012) and, at the same time, it carried a bacteriophage conveying the genes encoding the Shiga toxin type 2 subtype a (*stx2a*) (Bielaszewska et al., 2011; Scheutz et al., 2011).

The impact of the German outbreak was so deep that the causative STEC strain became iconic of a new *E. coli* pathotype: the Enteroaggregative Hemorrhagic *E. coli* (EAHEC) (Brzuszkiewicz et al., 2011). It has to be noted that *E. coli* strains matching the virulence genes profile of the EAHEC O104:H4 had been previously observed in a few occasions. At the beginning of the 90s, a small HUS outbreak occurred in France (Morabito et al., 1998). The episode was associated to infection with an *E. coli*

O111:H2 strain possessing the ability to colonize the host gut by the staked-brick adhesion mechanism typical of EAggEC (Nataro and Kaper, 1998) but which was also able to elaborate Stx2 (Morabito et al., 1998). Additionally, after the German outbreak caused by EAHEC O104:H4, a few sporadic cases of infection with similar EAHEC strains were retrospectively described in the time period 2000–2010 (Iyoda et al., 2000; Scavia et al., 2011). Finally, an HUS case, associated with EAHEC O111:H21 and an outbreak of EAHEC O127:H4 infections occurred, in Northern Ireland in 2012 (Dallman et al., 2012) and in Italy in 2013 (unpublished), respectively.

Up to the present day, four different EAHEC serotypes have been therefore identified including the O111:H2 (Morabito et al., 1998), O104:H4 (Bielaszewska et al., 2011), O111:H21 (Dallman et al., 2012), and O127:H4. This observation, together with the reported isolation of *Enterobacteriaceae* other than *E. coli* producing Stx from cases of human disease (Tschape et al., 1995; Paton and Paton, 1996) suggests that *stx*-phages can infect a range of bacterial hosts wider than expected. Further evidences supporting this hypothesis had been provided by Schmidt and co-workers, who were able to infect and lysogenize different *E. coli* pathogroups including EPEC, STEC, EAggEC, and EIEC using a chloramphenicol-resistant derivative of an Stx2-encoding Bacteriophage (Schmidt et al., 1999).

Recently the susceptibility of EAggEC strains of different serotypes to infection with the stx2-phage P13374 obtained from the German EAHEC O104:H4 strain has been investigated (Beutin et al., 2012). In that study, however, all the strains tested were found to be resistant to a high infective dose of P13374 and the authors concluded that the phage used had a restricted host range among EAggEC.

In order to contribute additional evidences to the subject matter, we investigated the ability of a panel of six *stx2*-phages to infect and lysogenize *E. coli* strains belonging to all the know pathogroups, including both diarrheagenic (DEC) and extraintestinal pathogenic *E. coli* (ExPEC). We describe that, at least in the conditions used in the laboratory, all the phages used were able to infect some of the strains belonging to all the *E. coli* pathogroups. Additionally, one of the phages used was able to produce lysogens visible after one cultural passage, which were no longer observable at the following passage.

Our observations suggest that the *stx*-phages used have a broad host specificity toward different *E. coli* pathogroups but that their stable acquisition might be a rare event.

MATERIALS AND METHODS

BACTERIAL STRAINS AND *stx2*-PHAGES

All the strains used in the present study are part of the collections held at Istituto Superiore di Sanità. A total of 33 *E. coli* strains were used as recipients for infection experiments. In detail, five typical Enteropathogenic *E. coli* (tEPEC), five atypical EPEC strains (aEPEC), five Enteroaggregative *E. coli* (EAggEC), five Enterotoxigenic *E. coli* (ETEC), three Enteroinvasive *E. coli* (EIEC), five Extraintestinal Pathogenic *E. coli* (ExPEC) strains isolated from urinary tract infections and five non-pathogenic *E. coli* strains from the ECOR collection (Ochman and Selander, 1984) were infected with *stx2*-phages. Typical and atypical EPEC were positive to PCR specific for the intimin-coding *eae* gene (Oswald et al., 2000). All typical EPEC also possessed the EAF plasmid as assessed by PCR of the BfP-coding gene (Franke et al., 1994).

All the EAggEC possessed the *aat* (Schmidt et al., 1995), *aggR* and *aaiC* (Boisen et al., 2012) genes as assessed by PCR in the conditions described in the respective papers.

PCR amplification was also used to identify the presence of the heat-stable (ST) and heat-labile (LT) enterotoxins coding-genes (Liu et al., 2013), *ipaH* (Liu et al., 2013) and the genes encoding the cytotoxic-necrotizing factor (CNF) (Kadhum et al., 2006) to verify the pathogroups ETEC, EIEC and ExPEC, respectively, in the conditions indicated in the respective papers.

E. coli K12 strains LE392 and DH5α have been used as control strains in infection experiments and for the *stx2*-phages amplification.

The *E. coli* strains CB553/5 and C125-06 (courtesy of Dr. Flemming Scheutz) and ED 191, ED 924, ED 703 and ED 508, were used to obtain the *stx2*-phages used for the infections by the spot agar assays. The CB553/5 strain is an ETEC O166:H15 possessing both *stx2* and LT genes and isolated from a human case of non-complicated diarrhea in Denmark. Strain C125-06 is a STEC O103:H25 isolated during a large HUS outbreak occurred in Norway in 2006 (Schimmer et al., 2008), possessing an *stx2*-phage with the gene encoding the Stx2 B subunit interrupted by an IS*1203* (this study, data not shown), which doesn't produce an active Shiga toxin. Strains ED 191 (O111:H2) and ED 924 (O127:H4) are Enteroaggregative-Hemorrhagic *E. coli* (EAHEC) strains isolated from small HUS outbreaks occurred in France in 1992 (Morabito et al., 1998) and in Northern Italy in 2013 (unpublished), respectively. Strain ED 703 is an O104:H4 EAHEC isolated from a HUS case occurred in Italy in 2009 (Scavia et al., 2011). Strain ED 508 is a STEC O157 producing Stx2 and isolated in Italy from a human case of disease. The *stx2*-phages derived from these *E. coli* strains have been termed with the strain's name preceded by the prefix Phi (e.g., Phi-508 from strain ED 508).

The panel of six Stx2-converting phages used to set up the infection experiments was selected using as a criterion their association with STEC of public health relevance. As a matter of fact, all the *stx*-phages conveyed the Stx2-coding genes, the toxin type associated with the most severe forms of STEC infections. Additionally, phages conveying such Stx type are the only ones so far associated with *E. coli* pathogroups other than the classical STEC.

PHAGES INDUCTION AND AMPLIFICATION

The *E. coli* strains carrying the *stx2*-phages were exposed to UV light in order to induce the excision of the phage genome from the bacterial chromosome (Sambrook and Russell, 2001). In detail, each bacterial strain was grown in Luria-Bertrani (LB) broth (Oxoid Limited, Basingstoke Hampshire, UK) overnight at 37°C with vigorous shaking. The culture was diluted 1:100 in LB modified broth (LB with 0.001% thiamine V/V) grown to 0.5 OD 600 and centrifuged. The bacterial pellets were re-suspended in a sterile solution of $CaCl_2$ 10 mM. The culture was exposed to UV light (130 μ Joule × 100) in a "Stratalinker® UV crosslinker" (Stratagene Cloning Systems, La Jolla, CA, USA). After induction, the culture was diluted in LB modified broth and incubated

at 37°C for 5 h with vigorous shaking. The culture was centrifuged and the supernatant containing phages particles filtered with 0.45 μm pore-filters. 100 μl of phage particles suspension were added to 100 μl of a culture of the propagator strain *E. coli* LE392 grown in LB modified broth at 0.5 OD 600 and the resulting mixture was maintained at 37°C for 20 min with static incubation. Each tube was added with 3.5 ml of LB modified soft agar (LB modified broth with agar 7 g/L) kept at 42°C and immediately poured on LB modified agar plates (LB modified broth with agar 15 g/L). Plates were incubated overnight at 37°C.

Four ml of SM buffer (100 mM NaCl, 8 mM $MgSO_4 \cdot 7H_2O$, 50 mM Tris-HCl 1 M pH 7.5, Gelatin 0.002%) were dispensed to each plate in order to recover phages particles from the lytic plaques and kept overnight at 4°C. The phage suspension in SM was recovered and added with chloroform at 5% final concentration. The phage suspension has been centrifuged at 500× g 10 min twice for removing agar debris and used to re-infect the propagator *E. coli* strain LE392 in the conditions described above in order to increase the phage titre. Finally, the phage suspension was concentrated by using Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-30 tubes (Merck Millipore, Billerica, MA, USA) with a cut-off of 30 KDa. Final phage titres ranged from 1×10^{10} to 4.9×10^{11} PFU/ml.

DETERMINATION OF THE *stx2*-PHAGES INTEGRATION SITES IN THE *E. COLI* GENOME

The *stx2*-phages integration sites in the *E. coli* strains were determined. The occupancy of loci *sbcB*, *wrbA*, *yehV*, and Z2577 was assessed as previously described (Serra-Moreno et al., 2007). An additional primer pair (YecE_fwd GCTAGCGCCGAGCAGC ACAA/YecE_rev ATGGCCGATGGCACCTGTCT) was deployed for specifically investigating the integrity of the locus *yecE* (this study).

SHIGATOXIN-GENES SUBTYPING

The determination of *stx*-genes subtypes was performed by PCR as previously described (Scheutz et al., 2012).

PLAQUE LIFT AND HYBRIDIZATION

Plaque blot experiments were performed to assay the homogeneity of the phages suspensions (Sambrook and Russell, 2001). In detail, phage induction experiments were performed as described above. A nylon membrane (Hybond N^+, GE Healthcare Life Sciences, UK) was placed onto the surface of each agar plate containing the phages plaques originated by lysis of the propagator strain and kept in place for 5 min. The filters were lifted from the plates and incubated with denaturation buffer (1.5 M NaCl, 0.5 M NaOH) at room temperature for 5 min and dried on a Whatman® paper 3 mm (GE Healthcare Life Sciences, UK). Filters were transferred to a new Whatman® paper 3 mm sheet impregnated with neutralization buffer (0.5 M Tris-Cl pH 7.2, 1 M NaCl) for 5 min. The membranes were dried and incubated with SSC 2× (6 M NaCl and 0.6 M sodium citrate) for 5 min. Once dried, 1 μl of a positive control, made up by unlabeled DNA corresponding to the probe used in the following hybridizations, was spotted on each filter. DNA was fixed by UV light (1200 μ Joule × 100) in a crosslinker "Stratalinker® UV crosslinker" (Stratagene Cloning Systems, La Jolla, CA, USA).

A DNA fragment obtained by PCR amplification of the *stx2*-gene using the primers Stx2F/Stx2R (Paton and Paton, 2002) was labeled by incorporating the digoxygenin-11-deoxyuridine-triphosphate using the PCR DIG probe synthesis Kit (Roche Diagnostics, Switzerland) and used as a probe. A prehybridization step was performed incubating the membranes in a buffer containing SSC 5×, 0.01% N-Lauryl Sarcosine, 1% blocking reagent (Roche Diagnostics, Switzerland) at 68°C for 1 h. The hybridization was carried out by incubating the filters overnight with 500 ng of the labeled probe at 68°C. The membranes were washed in SSC 2× with 0.1% SDS two times with agitation for 5 min followed by two washes in SSC 0.04× with 0.1% SDS at 64°C 15 min. The detection of the hybridized Digoxygenin-labeled probe was carried out with Anti- Digoxigenin AP Fab fragments and the detection reagent NBT/BCIP solution following the manufacturer's instructions (Roche Diagnostics, Switzerland).

INFECTION EXPERIMENTS BY SPOT AGAR ASSAY

The susceptibility of *E. coli* strains to the *stx2*-phages was assessed by spot agar assay as described elsewhere (Muniesa et al., 2004). In detail, each host strain was grown in LB medium overnight with vigorous shaking, diluted 100 times in LB modified broth (LB with 0.001% thiamine V/V) and grown at 0.5 OD 600. One hundred μl of the culture were added with 4 ml of LB modified soft agar (LB modified broth with agar 7 g/L) at 42°C and immediately poured on LB modified agar plates (LB modified broth with 15 g/L agar). Ten μl of each *stx2*-phage was spotted on each recipient strain immediately after the solidification of the soft agar layer. Phage titres used were: 1×10^{10} PFU/ml for the *stx2*-phages Phi- C125-06, Phi-703 and Phi-191, 1.5×10^{10} PFU/ml for Phi-CB553/5, 1.9×10^{10} PFU/ml for Phi-508 and 4.9×10^{11} for the *stx2*-phage from the EAHEC ED 924, Phi-924. Plates were incubated overnight at 37°C.

ASSESSMENT OF TOXIN PRODUCTION AND COLONY BLOT

In all the cases where lytic areas were not visible, the infection of the recipient strains was verified by carrying out Vero cell assays for the identification of the Stx production (Caprioli et al., 1994).

In detail, part of the top agar taken from the plates in correspondence with the area where the *stx*2-phage suspension was applied during the spot agar assay, was incubated in 2 ml of Tripticase Soya Broth (TSB, Oxoid Limited, Basingstoke Hampshire, UK) overnight at 37°C. One ml of the culture was centrifuged at 13000×g 10 min and the supernatant was filtered with 0.45 μm pore-filters. Twenty μl of each supernatant were inoculated on a semi-confluent monolayer of Vero cells. The presence of cythopathic effect was assessed at 24 h and confirmed at 48 and 72 h.

The *E. coli* recipients infected with the *stx*2-phage Phi-C125-06:IS*1203*, which does not induce the production of an active Stx, were verified by colony blot. Scalar dilutions of the overnight cultures obtained as described above were diluted and plated on LB agar plates in way of having about 50 colonies on each of the plates.

Bacterial colonies were transferred to a nylon membrane with high affinity for nucleic acids (Hybond N^+, GE Healthcare Life

Sciences, UK) by keeping the filter in contact with the surface of the agar for 5 min. The filters were incubated with denaturation buffer (1.5 M NaCl, 0.5 M NaOH) for 20 min and dried on a Whatman® paper 3 mm (GE Healthcare Life Sciences, UK). A further incubation on a Whatman® paper 3 mm impregnated with a buffer containing Triton X-100 0.2 % and NaOH 0.05 N was carried out for 10 min. Once dried, each filter was moved to a new sheet of Whatman® paper 3 mm impregnated with neutralization buffer (0.5 M Tris-Cl pH 7.2, 1 M NaCl) and dried again. The hybridization and detection steps were carried out as described above in the section "Plaque Lift and Hybridization."

RESULTS

CHARACTERIZATION OF THE *stx2*-PHAGES USED IN THE INFECTION EXPERIMENTS

Although all the phages used in the infection experiments were obtained from epidemiologically unrelated *E. coli* strains with different genomic background, including STEC, EAggEC and ETEC, they have been partially characterized to ascertain that we were not dealing with a homogeneous phage population made up by identical phage types.

Recently a peculiar type of *stx2*-phage has been described as being associated with STEC O157 isolated from human infections (Tozzoli et al., 2014). Such a phage, termed Φ8, presented large genomic differences when compared to the sequence of the reference *stx2*-phage BP933W, present in the reference STEC O157 strain EDL933, in the regions where the Stx-coding genes are located and containing the genes regulating the switch between lysogeny and lytic cycle, respectively (Tozzoli et al., 2014). All the phages in the panel used in the infection experiments were negative to a Φ8-specific PCR (Tozzoli et al., 2014) but the Phi-508 phage, indicating that only the latter belonged to this particular phage type.

The whole genome sequences (WGS) of the strains CB553/5, ED 508, ED 703 and ED 924 and of the phage Phi-191 from the EAHEC O111:H2 that caused the outbreak in France in the 90s have been determined (data not shown). The contigs composing the WGS of the bacterial strains were mapped against the genomic sequence of the phage Phi-191 (data not shown).

The comparison showed that in the genomes of the three EAHEC strains were present stx2-phages showing an average similarity with the Phi-191 of 90% and higher, indicating that the three phages belonged to a homogeneous population. Conversely, the contigs of the ETEC strain CB553/5 only mapped in a small region of 7750 bp containing the genes encoding the Stx2, indicating that, in this strain the *stx2*-phage must be extensively different from that present in the EAHEC strains.

The sequence of the phage the phage Phi-C125-06 corresponds to the published sequence of phage TL-2011c, isolated from a STEC O103:H25 during an outbreak of HUS cases occurred in Norway in 2006 (Schimmer et al., 2008). We found that the sequence of phage Phi-191 was about 90% homologous to that of TL-2011c (data not shown), similarly to how reported for the phage P13374, present in the EAHEC O104:H4 that caused the German outbreak in 2011 (Beutin et al., 2012). Additionally, in the Phi-C125-06 an IS 1203 was detected in the gene encoding the B subunit of the Stx.

Finally, we observed that the contigs composing the WGS of the strain ED 508 that mapped on the sequence of the phage Phi-191 shared an average 65% of homology at nucleotide level.

Taken together these findings indicate that the stx2-phages used in the infection experiments belonged to at least three different types.

The PCR sub-typing scheme deployed by Scheutz and coworkers was used to identify the subtypes of the Shiga toxin-coding genes harbored by the *stx2*-phages used in this study (Scheutz et al., 2012). All the phages harbored $stx2_a$ subtype, with the exception of Phi-CB553/5, which possessed $stx2_d$.

The insertion sites in the parental bacterial strains were determined for all the *stx2*-phages but the Phi-C125-06, which was already described to be present in the *wrbA* locus (Sekse et al., 2008a). With the exception of Phi-508, which is localized in *yehV* locus, all the other phages composing the set used in this study seemed to be inserted in *wrbA*.

The homogeneity of the phage suspensions used in the experimental infections was determined by infecting the propagator K 12 strain LE 392 and by performing plaque lift and hybridization with a probe corresponding to the *stx2* gene encoding the Stx2 A subunit. All the plaques present on the plates were positive to hybridization with the *stx2* probe, indicating that the all the phage suspensions used in the following infection experiments contained a unique bacteriophage population.

SUSCEPTIBILITY OF *E. COLI* STRAINS TO THE INFECTION WITH THE *stx2*-PHAGES

E. coli strains belonging to all the known pathotypes were infected with the set of six *stx2*-phages. The results of the spot agar tests are reported in **Table 1**. In total, we assayed the susceptibility of a panel of 35 *E. coli* strains, including five typical EPEC (tEPEC, BfP-positive), five atypical EPEC (aEPEC, BfP-negative), five ETEC, five EAggEC strains, three EIEC, five ExPEC, five *E. coli* non-pathogenic strains belonging to the ECOR collection (Ochman and Selander, 1984) and two *E. coli* K 12 laboratory strains.

With the exception of the bacteriophage obtained from the STEC/ETEC strain (CB553/5), which induced visible lysis in all the strains infected, all the other phages were able to induce clear lysis halos only in *E. coli* strains belonging to the Diarrheagenic *E. coli* (DEC) pathogroups, although with different efficiency (**Table 1**), but not in ExPEC cultures. Finally, the two *E. coli* K 12 strains used as controls demonstrated to be susceptible to infection with all the phages tested, showing marked lytic areas.

These results suggested that the *stx2*-phages used in this study, with the exception of the Phi-CB553/5, might have a specificity for DEC.

To ascertain if the lack of lytic areas in the negative *E. coli* cultures was due to mechanisms of resistance to infection, further analyses were carried out aiming at determining the production of Stx. Such an approach was based on the assumption that, for the toxin to be produced, the phage DNA must be inserted into the chromosome of the bacterial host and the host itself must therefore be susceptible to infection, also in the absence of visible lysis. The occurrence of infection in the absence of lytic areas has been previously reported during experimental infection of a panel of

Table 1 | Results of the spot agar assays showing the susceptibility of the *E. coli* strains to infection with the different *stx2*-phages used.

Pathogroup (No. of strains)	No. of strains showing lysis upon infection with Stx2-Phages					
	Phi-C125-06	Phi-703	Phi-191	Phi-508	Phi-924	Phi-CB553/5
EIEC (3)	1	3	3	3	2	3
ExPEC (5)	0	0	0	0	0	5
tEPEC (5)	0	1	1	0	1	5
aEPEC (5)	2	0	3	1	1	5
ETEC (5)	0	1	1	5	0	5
EAggEC (5)	1	0	2	1	1	5
ECOR collection (5)	1	3	4	2	2	5
E. coli K 12 (2)	2	2	2	2	2	2
Tot. (35)	7	10	16	14	9	35

diarrheagenic *E. coli* with a derivative chloramphenicol-resistant *stx2*-phage (Schmidt et al., 1999).

All the bacterial strains used in the experimental infections have been assayed for the capability to induce a cythopathic effect (CPE) onto Vero cells monolayers before infection with the *stx2*-phages and most of them proved negative by microscopic observation of the cell up to 72 h post inoculum. Of the 35 *E. coli* strains assayed two EXPEC, one tEPEC, two aEPEC, one EAggEC, and four ETEC strains induced the death of Vero cells and have therefore been excluded from the following screening by Vero cells assay (VCA).

The bacterial strains that gave negative results in the VCA and that, at the same time, did not show lysis in the spot agar assay, were tested again after the experimental infection with the phages by inoculating the supernatant of bacterial cultures obtained by incubating in a liquid medium the area of the plates where the phage suspension was applied on the test *E. coli* strains, onto Vero cell monolayers.

By using this approach all the cultures showed a cytopathic effect (CPE). The CPE observed appeared after 24 h incubation and progressed toward the complete death of the monolayer in 48–72 h.

An alternative strategy was adopted to verify if the lack of lytic areas in the test strains infected with phage Phi-C125-06 was due to mechanisms of resistance. As a matter of fact, this phage doesn't possess an intact Stx2 B subunit-coding gene and therefore would not induce any CPE on Vero cells monolayers. The bacterial cultures derived by inoculating the area of the spot agar test where the phage was applied were plated onto solid media and subjected to colony blot using a probe corresponding to a region of the gene encoding the Stx2 A subunit. None of the colonies were positive in the hybridization experiments, suggesting that the strains had not been infected.

VERIFICATION OF THE STABLE ACQUISITION OF *stx2*-PHAGES BY THE *E. COLI* STRAINS

Following the observation that *E. coli* strains belonging to all the pathogroups used were susceptible to infection with the *stx2*-phages used, we investigated on the possibility that the *E. coli* strains were able to maintain the phage DNA stably integrated in their chromosome. Sub-cultures of the infected strains that produced Stx were diluted and analyzed by colony blot to verify the presence of the *stx2* gene in their genome. Lysogens were only observed in the cultures of one ETEC and one EAggEC strains as well as in the *E. coli* K 12 strain LE 392 used as control, all infected with the Phi 191 phage. The latter lysogen was stable and could be cultured several times, while the cultures from the pathogenic *E. coli* became negative already at the second cultural passages after the phage infections.

DISCUSSION

STEC are human pathogens whose complex nature and pathogenicity mechanisms are not completely understood yet. As a matter of fact, a number of different geno-phenotypes have been described so far in STEC strains isolated from human infections displaying a wide and diverse range of symptoms including mild diarrheal disease as well as life threatening forms such as the HUS (Tozzoli and Scheutz, 2014). The association between the production of Stx and the ability to cause the A/E lesion, has been long regarded as the virulence features asset characterizing the STEC causing HUS (Levine, 1987), with the understood feeling that the virulence gene repertoires described in other STEC had to be associated with the less severe forms of the infection. This approach led to the definition of schemes attempting at categorizing STEC for the purpose of laying the ground for a proper epidemiological approach and management of infections. The most comprehensive among those schemes was developed by Karmali and co-workers a decade ago (Karmali et al., 2003) and considered either the clinical aspects of the infection or the virulence features of the STEC strains. Such a scheme used the term seropathotype (SPT) to define the different groups of STEC, with the SPTs A and B including those causing the most severe forms of infection or associated with outbreaks (Karmali et al., 2003). Notably, the STEC included in the SPTs A and B are all capable to cause the A/E lesion.

This reference scheme has been efficaciously used to frame the STEC isolated from human disease for many years. Nevertheless, in 2011, one of the largest and most severe outbreaks of STEC infection occurred in Germany and France (Frank et al., 2011; Mariani-Kurkdjian et al., 2011) and caused the paradigm to vault in a new direction. The infecting strain, a STEC O104:H4 was undoubtedly associated with HUS, with an impressive proportion

of cases of infection progressing toward this severe syndrome (Frank et al., 2011) and, at the same time, it was not able to induce the A/E lesion (Bielaszewska et al., 2011).

The investigation on the genetic asset of the outbreak strain revealed that beside the ability to produce Stx, it possessed the virulence genes encoding the adhesion machinery of Enteroaggregative *E. coli* (EAggEC) (Bielaszewska et al., 2011) and it has been proposed that this strain belongs to a new pathogroup of *E. coli* termed Enteroaggeregative Haemorrhagic *E. coli* (EAHEC) (Brzuszkiewicz et al., 2011). Several hypothesis have been made on the evolution of the EAHEC O104:H4 including the opposite views of the derivation of this strain from an EAggEC that acquired an *stx2*-phage by horizontal gene transfer (Brzuszkiewicz et al., 2011; Rasko et al., 2011) or proposing its evolution from an ancestor STEC O104:H4 by stepwise gain and loss of chromosomal and plasmid-encoded virulence factors (Mellmann et al., 2011). More recently evidences have been provided showing that the EAHEC O104:H4 could have evolved by the uptake of a *stx2*-phage originated from the bovine reservoir by an EAggEC O104:H4 (Beutin et al., 2013). This scenario parallels what it can be inferred for the STEC strains causing the attaching and effacing lesion, such as those belonging to SPTs A and B (Karmali et al., 2003) which are also termed Enterohemorrhagic *E. coli* (EHEC) (Levine, 1987). As a matter of fact, the latter can be considered as atypical EPEC (aEPEC) that developed the capability to produce Stx (Trabulsi et al., 2002). As in the case of EAggEC and EAHEC, aEPEC and EHEC share the same mechanism of colonization and are mainly distinguished by the capability to produce Stx, which increase their pathogenicity.

The above considerations bring into question if STEC are indeed an *E. coli* pathogroup in its own right, or rather if this pathogen represents multiple pathogroups whose virulence has been increased by the event of an *stx*-phage acquisition. Or even if the phage itself might be considered as being the pathogen, using the *E. coli* colonization machinery to establish a successful infection in the final host. Whatever scenario is chosen, it is undeniable that for any intestinal infection to be successful, a pathogen must efficiently colonize the mucosa, overcoming the competition with the resident microflora. In this respect, both the SPTs A and B STEC and EAHEC can rely on efficient colonization machineries, with the one present in the latter being recognized as causing the most diffuse and long-lasting intestinal colonization (Nataro and Kaper, 1998). Accordingly, while both the groups have been associated with HUS, the EAHEC O104:H4 infections during the German outbreak of 2011 progressed toward this severe form in about 30% of cases (Frank et al., 2011), against the typical 5%–10% of cases of infection with SPT A and B STEC strains developing HUS (Tarr et al., 2005).

How stated above admits the possibility that any *E. coli* strain could acquire the ability to produce Stx and that, virtually any pathogenic *E. coli* might stem a stable augmented-pathogenicity variant threatening the public health systems. This is hold true for EAggEC, which beside the German outbreak have been described as being associated with HUS at least in other two outbreaks occurred in the 90s in France (Morabito et al., 1998) and in 2013 in Italy (unpublished) as well as in some unrelated cases recorded in Europe and Asia in the period 2000–2012 (Iyoda et al., 2000; Scavia et al., 2011; Beutin et al., 2012; Dallman et al., 2012).

In the conditions used in our laboratory we observed that, beside the EAggEC, other *E. coli* pathogroups were susceptible to infection with different *stx2*-phages obtained from a number of different STEC types including EAHEC, typical STEC O157 and non-O157 as well as from an hybrid ETEC/STEC strain isolated from a patient with diarrhea in Denmark (Dr. F. Scheutz, personal communication). All the phages assayed conveyed the genes encoding the Stx2 type (predominantly Stx2a subtype), the one associated with STEC causing HUS and the only one found in Stx-producing EAggEC so far.

We ascertained that in the conditions used in the laboratory, all the DEC types assayed, including tEPEC, aEPEC, EAggEC, ETEC, EIEC, were susceptible to infection with the *stx2*-phages. Additionally, the *stx2*-phage from the hybrid ETEC/STEC strain CB553/5 infected all the *E. coli* strains tested, including ExPEC, and the field isolates of non-pathogenic *E. coli* of the ECOR collection.

We observed that the *stx2*-phages used in this study seem not to have specificity for particular *E. coli* groups although the efficacy of the infections varied with the phage used (**Table 1**).

Our results are in line with the identification of the Stx2-producing ETEC described above, and with the recent report of a case of septicemia in a human patient with evidence of infection with an *E. coli* strain possessing the genes encoding the Stx2 and matching the genetic background of an ExPEC (Wester et al., 2013). Our results are also in agreement with how reported in a previous study on the transduction of a chloramphenicol-resistant *stx2*-phage where *E. coli* strains belonging to all the diarrheagenic *E. coli* groups were successfully infected and produced stable lysogens (Schmidt et al., 1999). These findings provide evidences that the *stx2*-phage acquisition can involve a spectrum of *E. coli* hosts wider than expected and that no pathogroup-specific barriers seem to exist to the acquisition of *stx2*-phages belonging to the types used in this work.

However, it cannot be excluded that *stx*-phages different from those used in this study might have a restricted host range and that in the natural environment the spreading of the *stx*-phages to *E. coli* hosts might be hindered by factors related with the bacterial hosts. The existence of such restrictions could explain how reported in a recent study where the susceptibility of a panel of 31 EAggEC to infection with the *stx2*-phage derived from the EAHEC O104:H4 that caused the German outbreak in 2011 was assessed (Beutin et al., 2012). The authors observed that none of the bacterial strains tested showed evidences of infection with the *stx2*-phage in spite of the high titer used (Beutin et al., 2012).

Differences in the host range associated with the phage type could be also hypothesized by considering the results of the spot agar tests performed in this study. As a matter of fact, we observed that while the phage Phi-CB553/5 induced clear lysis in all the strains analyzed, all the other phages infected a variable number of strains in each of the pathogroups but never induced lysis in the ExPEC strains analyzed. Nevertheless we could observe that such strains induced CPE onto Vero cells monolayers after the infection with the *stx2*-phages. This observation, although not confirmed by sero-neutralization, suggests a wider host range of the phages

used in this study. Further work is needed to clarify the mechanisms leading to the selection of the *stx*-phages able to infect the different *E. coli* pathogroups.

Interestingly, all the *E. coli* pathotypes, with the exception of typical STEC, have a human reservoir and an inter-human transmission cycle (Nataro and Kaper, 1998), while the typical STEC and the related *stx*-phages seem to have their reservoir in the intestinal tract of ruminants (Caprioli et al., 2005; Beutin et al., 2013). This observation implies that for the event of an *stx*-phage acquisition by a human *E. coli* to occur a common ecosystem must exist where the two organisms may encounter and interact. Such an ecosystem could be represented by the intestine of a mammalian host or the environment (Schmidt et al., 1999; Sekse et al., 2008b; Imamovic et al., 2009; Dopfer et al., 2010). Given the inter-human circulation of the EAggEC, the possibility that EAHEC may have arisen following an environmentally mediated event of an *stx2*-phage acquisition is interesting. As a matter of fact in low-income countries, where infections with DEC are endemic the ineffective treatments of human sewages may account for a wide dispersion of these pathogens in the environment, with the consequent possibility for them to come into contact with *stx*-phages originating from ruminant's excreta. Accordingly, an origin in these countries has been traced or postulated for some of the Stx2-producing EAggEC (Scavia et al., 2011; Beutin et al., 2012; Weiser et al., 2013), strengthening such a hypothesis for their derivation.

AUTHOR CONTRIBUTIONS

Rosangela Tozzoli conceived the experimental design and drafted the manuscript, Laura Grande carried out the experimental work and analysis, and revised critically the manuscript, Paola Ranieri carried out part of experiments (stx subtyping and insertion sites identification), Valeria Michelacci and Antonella Maugliani participated in the revision of the manuscript and supported Laura Grande in the experimental work, Alfredo Caprioli contributed to critical the revision of the draft manuscript for important intellectual content, Stefano Morabito conceived the study and strongly contributed to revise the manuscript. Finally, all the authors approved the manuscript to be published.

ACKNOWLEDGMENT

We would like to thank Dr. Flemming Scheutz for providing the strains CB553/5 and C125-06 and our colleagues Maria Luisa Marziano and Fabio Minelli for their valuable help in performing the Vero Cell Assay.

REFERENCES

Beutin, L., Hammerl, J. A., Reetz, J., and Strauch, E. (2013). Shiga toxin-producing *Escherichia coli* strains from cattle as a source of the Stx2a bacteriophages present in enteroaggregative *Escherichia coli* O104:H4 strains. *Int. J. Med. Microbiol.* 303, 595–602. doi: 10.1016/j.ijmm.2013.08.001

Beutin, L., Hammerl, J. A., Strauch, E., Reetz, J., Dieckmann, R., Kelner-Burgos, Y., et al. (2012). Spread of a distinct Stx2-encoding phage prototype among *Escherichia coli* O104:H4 strains from outbreaks in Germany, Norway, and Georgia. *J. Virol.* 86, 10444–10455. doi: 10.1128/JVI.00986-12

Bielaszewska, M., Mellmann, A., Zhang, W., Kock, R., Fruth, A., Bauwens, A., et al. (2011). Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. *Lancet Infect. Dis.* 11, 671–676. doi: 10.1016/S1473-3099(11)70165-7

Boisen, N., Scheutz, F., Rasko, D. A., Redman, J. C., Persson, S., Simon, J., et al. (2012). Genomic characterization of enteroaggregative *Escherichia coli* from children in Mali. *J. Infect. Dis.* 205, 431–444. doi: 10.1093/infdis/jir757

Boisen, N., Struve, C., Scheutz, F., Krogfelt, K. A., and Nataro, J. P. (2008). New adhesin of enteroaggregative *Escherichia coli* related to the Afa/Dr/AAF family. *Infect. Immun.* 76, 3281–3292. doi: 10.1128/IAI.01646-07

Brzuszkiewicz, E., Thurmer, A., Schuldes, J., Leimbach, A., Liesegang, H., Meyer, F. D., 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, 883–891. doi: 10.1007/s00203-011-0725-6

Caprioli, A., Luzzi, I., Rosmini, F., Resti, C., Edefonti, A., Perfumo, F., et al. (1994). Community-wide outbreak of hemolytic-uremic syndrome associated with non-O157 verocytotoxin-producing *Escherichia coli*. *J. Infect. Dis.* 169, 208–211. doi: 10.1093/infdis/169.1.208

Caprioli, A., Morabito, S., Brugere, H., and Oswald, E. (2005). Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission. *Vet. Res.* 36, 289–311. doi: 10.1051/vetres:2005002

Dallman, T., Smith, G. P., O'brien, B., Chattaway, M. A., Finlay, D., Grant, K. A., et al. (2012). Characterization of a verocytotoxin-producing enteroaggregative *Escherichia coli* serogroup O111:H21 strain associated with a household outbreak in Northern Ireland. *J. Clin. Microbiol.* 50, 4116–4119. doi: 10.1128/JCM.02047-12

Dopfer, D., Sekse, C., Beutin, L., Solheim, H., Van Der Wal, F. J., De Boer, A., et al. (2010). Pathogenic potential and horizontal gene transfer in ovine gastrointestinal *Escherichia coli*. *J. Appl. Microbiol.* 108, 1552–1562. doi: 10.1111/j.1365-2672.2009.04575.x

Frank, C., Werber, D., Cramer, J. P., Askar, M., Faber, M., An Der Heiden, M., et al. (2011). Epidemic profile of Shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany. *N. Engl. J. Med.* 365, 1771–1780. doi: 10.1056/NEJMoa1106483

Franke, J., Franke, S., Schmidt, H., Schwarzkopf, A., Wieler, L. H., Baljer, G., et al. (1994). 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.* 32, 2460–2463.

Imamovic, L., Jofre, J., Schmidt, H., Serra-Moreno, R., and Muniesa, M. (2009). Phage-mediated Shiga toxin 2 gene transfer in food and water. *Appl. Environ. Microbiol.* 75, 1764–1768. doi: 10.1128/AEM.02273-08

Iyoda, S., Tamura, K., Itoh, K., Izumiya, H., Ueno, N., Nagata, K., et al. (2000). Inducible stx2 phages are lysogenized in the enteroaggregative and other phenotypic *Escherichia coli* O86:HNM isolated from patients. *FEMS Microbiol. Lett.* 191, 7–10. doi: 10.1111/j.1574-6968.2000.tb09311.x

Kadhum, H. J., Ball, H. J., Oswald, E., and Rowe, M. T. (2006). Characteristics of cytotoxic necrotizing factor and cytolethal distending toxin producing *Escherichia coli* strains isolated from meat samples in Northern Ireland. *Food Microbiol.* 23, 491–497. doi: 10.1016/j.fm.2005.07.003

Karmali, M. A. (2009). Host and pathogen determinants of verocytotoxin-producing *Escherichia coli*-associated hemolytic uremic syndrome. *Kidney Int. Suppl.* S4–S7. doi: 10.1038/ki.2008.608

Karmali, M. A., Mascarenhas, M., Shen, S., Ziebell, K., Johnson, S., Reid-Smith, R., 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, 4930–4940. doi: 10.1128/JCM.41.11.4930-4940.2003

Levine, M. M. (1987). *Escherichia coli* that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. *J. Infect. Dis.* 155, 377–389. doi: 10.1093/infdis/155.3.377

Liu, J., Gratz, J., Amour, C., Kibiki, G., Becker, S., Janaki, L., et al. (2013). A laboratory-developed TaqMan Array Card for simultaneous detection of 19 enteropathogens. *J. Clin. Microbiol.* 51, 472–480. doi: 10.1128/JCM.02658-12

Mariani-Kurkdjian, P., Bingen, E., Gault, G., Jourdan-Da Silva, N., and Weill, F. X. (2011). *Escherichia coli* O104:H4 south-west France, June 2011. *Lancet Infect. Dis.* 11, 732–733. doi: 10.1016/S1473-3099(11)70266-3

Mcdaniel, T. K., and Kaper, J. B. (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. doi: 10.1046/j.1365-2958.1997.2311591.x

Mellmann, A., Harmsen, D., Cummings, C. A., Zentz, E. B., Leopold, S. R., Rico, A., et al. (2011). Prospective genomic characterization of the German

enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. *PLoS One* 6: e22751. doi: 10.1371/journal.pone.0022751

Morabito, S., Karch, H., Mariani-Kurkdjian, P., Schmidt, H., Minelli, F., Bingen, E., et al. (1998). Enteroaggregative, Shiga toxin-producing *Escherichia coli* O111:H2 associated with an outbreak of hemolytic-uremic syndrome. *J. Clin. Microbiol.* 36, 840–842.

Muniesa, M., Serra-Moreno, R., and Jofre, J. (2004). Free Shiga toxin bacteriophages isolated from sewage showed diversity although the stx genes appeared conserved. *Environ. Microbiol.* 6, 716–725. doi: 10.1111/j.1462-2920.2004.00604.x

Nataro, J. P., and Kaper, J. B. (1998). Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* 11, 142–201.

O'brien, A. D., Newland, J. W., Miller, S. F., Holmes, R. K., Smith, H. W., and Formal, S. B. (1984). Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* 226, 694–696. doi: 10.1126/science.6387911

Ochman, H., and Selander, R. K. (1984). Standard reference strains of *Escherichia coli* from natural populations. *J. Bacteriol.* 157, 690–693.

Okuda, T., Tokuda, N., Numata, S., Ito, M., Ohta, M., Kawamura, K., et al. (2006). Targeted disruption of Gb3/CD77 synthase gene resulted in the complete deletion of globo-series glycosphingolipids and loss of sensitivity to verotoxins. *J. Biol. Chem.* 281, 10230–10235. doi: 10.1074/jbc.M600057200

Oswald, E., Schmidt, H., Morabito, S., Karch, H., Marches, O., and Caprioli, A. (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. doi: 10.1128/IAI.68.1.64-71.2000

Paton, A. W., and Paton, J. C. (1996). Enterobacter cloacae producing a Shiga-like toxin II-related cytotoxin associated with a case of hemolytic-uremic syndrome. *J. Clin. Microbiol.* 34, 463–465.

Paton, A. W., and Paton, J. C. (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. doi: 10.1128/JCM.40.1.271-274.2002

Rasko, D. A., Webster, D. R., Sahl, J. W., Bashir, A., Boisen, N., Scheutz, F., et al. (2011). Origins of the E. coli strain causing an outbreak of hemolytic-uremic syndrome in Germany. *N. Engl. J. Med.* 365, 709–717. doi: 10.1056/NEJMoa1106920

Sambrook, J., and Russell, D., (ed.). (2001). *Molecular Cloning: a Laboratory Manual*. New York, NY: Cold Spring Harbor Laboratory Press.

Scavia, G., Morabito, S., Tozzoli, R., Michelacci, V., Marziano, M. L., Minelli, F., et al. (2011). Similarity of Shiga toxin-producing *Escherichia coli* O104:H4 strains from Italy and Germany. *Emerg. Infect. Dis.* 17, 1957–1958. doi: 10.3201/eid1710.111072

Scheutz, F., Nielsen, E. M., Frimodt-Moller, J., Boisen, N., Morabito, S., Tozzoli, R., 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. Available online at: http://www.eurosurveillance.org/viewarticle.aspx?articleid=19889

Scheutz, F., Teel, L. D., Beutin, L., Pierard, D., Buvens, G., Karch, H., et al. (2012). Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *J. Clin. Microbiol.* 50, 2951–2963. doi: 10.1128/JCM.00860-12

Schimmer, B., Nygard, K., Eriksen, H. M., Lassen, J., Lindstedt, B. A., Brandal, L. T., et al. (2008). Outbreak of haemolytic uraemic syndrome in Norway caused by stx2-positive *Escherichia coli* O103:H25 traced to cured mutton sausages. *BMC Infect. Dis.* 8:41. doi: 10.1186/1471-2334-8-41

Schmidt, H., Bielaszewska, M., and Karch, H. (1999). Transduction of enteric *Escherichia coli* isolates with a derivative of Shiga toxin 2-encoding bacteriophage phi3538 isolated from *Escherichia coli* O157:H7. *Appl. Environ. Microbiol.* 65, 3855–3861.

Schmidt, H., Knop, C., Franke, S., Aleksic, S., Heesemann, J., and Karch, H. (1995). Development of PCR for screening of enteroaggregative *Escherichia coli*. *J. Clin. Microbiol.* 33, 701–705.

Sekse, C., Muniesa, M., and Wasteson, Y. (2008a). Conserved Stx2 phages from *Escherichia coli* O103:H25 isolated from patients suffering from hemolytic uremic syndrome. *Foodborne Pathog. Dis.* 5, 801–810. doi: 10.1089/fpd.2008.0130

Sekse, C., Solheim, H., Urdahl, A. M., and Wasteson, Y. (2008b). Is lack of susceptible recipients in the intestinal environment the limiting factor for transduction of Shiga toxin-encoding phages? *J. Appl. Microbiol.* 105, 1114–1120. doi: 10.1111/j.1365-2672.2008.03845.x

Serra-Moreno, R., Jofre, J., and Muniesa, M. (2007). Insertion site occupancy by stx2 bacteriophages depends on the locus availability of the host strain chromosome. *J. Bacteriol.* 189, 6645–6654. doi: 10.1128/JB.00466-07

Tarr, P. I., Gordon, C. A., and Chandler, W. L. (2005). Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. *Lancet* 365, 1073–1086. doi: 10.1016/S0140-6736(05)71144-2

Tozzoli, R., Grande, L., Michelacci, V., Fioravanti, R., Gally, D., Xu, X., et al. (2014). Identification and characterization of a peculiar vtx2-converting phage frequently present in Verocytotoxin-producing Escherichia coli (VTEC) O157 isolated from human infections. *Infect. Immun.* doi: 10.1128/IAI.01836-14. [Epub ahead of print].

Tozzoli, R., and Scheutz, F. (2014). "Diarrhoeagenic *Escherichia coli* infections in humans," in *Pathogenic Escherichia Coli, Molecular and Cellular Microbiology, 1st Edn.*, ed. M. Stefano (Norfolk: Caister Academic Press), 1–18.

Trabulsi, L. R., Keller, R., and Tardelli Gomes, T. A. (2002). Typical and atypical enteropathogenic *Escherichia coli*. *Emerg. Infect. Dis.* 8, 508–513. doi: 10.3201/eid0805.010385

Tschape, H., Prager, R., Streckel, W., Fruth, A., Tietze, E., and Bohme, G. (1995). Verotoxinogenic Citrobacter freundii associated with severe gastroenteritis and cases of haemolytic uraemic syndrome in a nursery school: green butter as the infection source. *Epidemiol. Infect.* 114, 441–450. doi: 10.1017/S0950268800052158

Weiser, A. A., Gross, S., Schielke, A., Wigger, J. F., Ernert, A., Adolphs, J., et al. (2013). Trace-back and trace-forward tools developed ad hoc and used during the STEC O104:H4 outbreak 2011 in Germany and generic concepts for future outbreak situations. *Foodborne Pathog. Dis.* 10, 263–269. doi: 10.1089/fpd.2012.1296

Wester, A. L., Brandal, L. T., and Dahle, U. R. (2013). Extraintestinal pathogenic *Escherichia coli* carrying the Shiga Toxin gene stx2. *J. Clin. Microbiol.* 51, 4279–4280. doi: 10.1128/JCM.01349-13

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 April 2014; accepted: 01 June 2014; published online: 20 June 2014.
*Citation: Tozzoli R, Grande L, Michelacci V, Ranieri P, Maugliani A, Caprioli A and Morabito S (2014) Shiga toxin-converting phages and the emergence of new pathogenic Escherichia coli: a world in motion. Front. Cell. Infect. Microbiol. **4**:80. doi: 10.3389/fcimb.2014.00080*
This article was submitted to the journal Frontiers in Cellular and Infection Microbiology.

CHAPTER 7

Discussion and Conclusions

E. coli is a versatile microorganism. It has been described as a commensal of mammals and other animals, part of their gut microflora, a harmless ubiquitous microorganism inhabiting a wide range of environmental niches, and at the same time the causative agent of diseases in human and animals sometimes causing large outbreaks with thousands cases of infections. In between such extreme circumstances, intermediate scenarios do exist where *E. coli* strains cause infections displaying a wide range of symptoms and spanning from extra-intestinal forms to gastrointestinal afflictions, the latter in turn ranging from mild diarrhoea to severe syndromes such as bloody diarrhoea, and the life-threatening haemolytic uremic syndrome (HUS). *E. coli* causing intestinal illnesses, termed Diarrheagenic *E. coli* (DEC), show the highest diversity. Verocytotoxin-producing *E. coli* (VTEC) are iconic of such a variety and are the DEC group causing the most severe forms of infections (See section 1.2.6 and Fig. 1.2). VTEC are important public health hazards in industrialized countries and are the only pathotype of *E. coli* with an ascertained animal reservoir, represented by the gastro-intestinal tract of ruminants (Caprioli *et al.*, 2005). The vehicles operating the transmission of VTEC infections to humans are diverse and principally include food or water primarily contaminated. Additionally, environmental exposure or direct contacts with animals may be involved in the transmission of the infections as well as the inter-human contacts (Heuvelink *et al.*, 2002).

VTEC pathogenicity mainly relies upon the production of Verocytotoxins (VTs), while the strains typically causing HUS have the ability to cause a typical histopathological lesion to the host enterocyte known as Attaching and Effacing (A/E). The latter is conferred by the presence of a pathogenicity island referred to as the LEE locus. VTEC isolates possessing this feature have been often designated as Enterohaemorrhagic *E. coli* (EHEC) (Caprioli *et al.*, 2005). Although the LEE pathogenicity island is considered a hallmark of the VTEC causing HUS, LEE-negative isolates belonging to different serogroups have been isolated from cases of human disease including HUS (Steyert *et al.*, 2012) bringing into question the size and the composition of the "virulome" of the VTEC strains pathogenic to humans. It has been proposed that different forces intervened in the evolution of the different VTEC types, since the gain and loss of virulence factors such as the LEE locus, the Enterohaemolysin-coding gene and the VT-converting phages themselves, have occurred several times and in different lineages (Franz *et al.*, 2014).

VTEC O157:H7 is paradigmatic of such a complex evolution. Some authors hypothesised that VTEC O157:H7 strains are genetically related to EPEC O55:H7 (Wick *et al.*, 2005) and proposed a model of step-wise emergence from a common ancestor LEE-positive strain able to cause the A/E lesion. In a first step, the acquisition of a VT-phage gave origin to a VTEC O55:H7, which later gained the pO157 plasmid and switched to O157 serogroup generating an intermediate clone. Two separate lineages diverged at that time, one non-motile sorbitol-fermenting VTEC O157:H- and the VTEC O157:H7 non sorbitol-fermenting (Fig. 7.1).

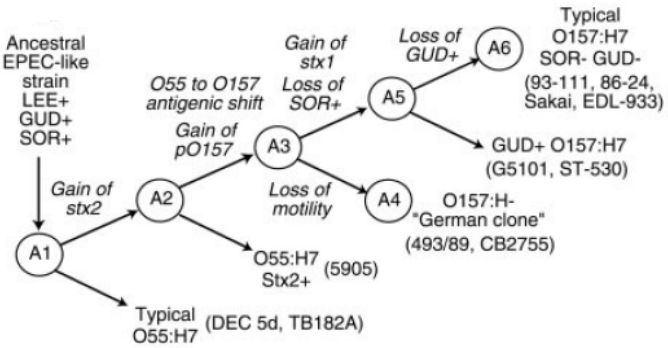


Fig 7.1: Evolutionary genomic changes in the emergence of E. coli O157:H7.
Stepwise model for the evolution of E. coli O157:H7 from an Enteropathogenic E. coli-like ancestor (Wick et al., 2005).

The latter may have undergone further diversification into two distinct lineages with different ecological characteristics as it has been proposed by other authors (Kim *et al.*, 1999).
The stepwise model has been also postulated for the evolution of non-O157 VTEC, which may have emerged from an unknown common ancestor by independent acquisition events of Mobile Genetic Elements (MGE)-encoded virulence factors (Coombes *et al.*, 2011).
The large diversity of VTEC may result from the existence of a pool of accessory genetic information in form of mobile genetic elements (MGEs) that has been made available for the exchanges occurred during the evolution of the different pathogenic *E. coli*. As a matter of fact all the *E. coli* virulence factors described so far are carried by MGEs that have been acquired and stabilized in the pangenome characterising the various *E. coli* pathotypes (Fig 7.2).

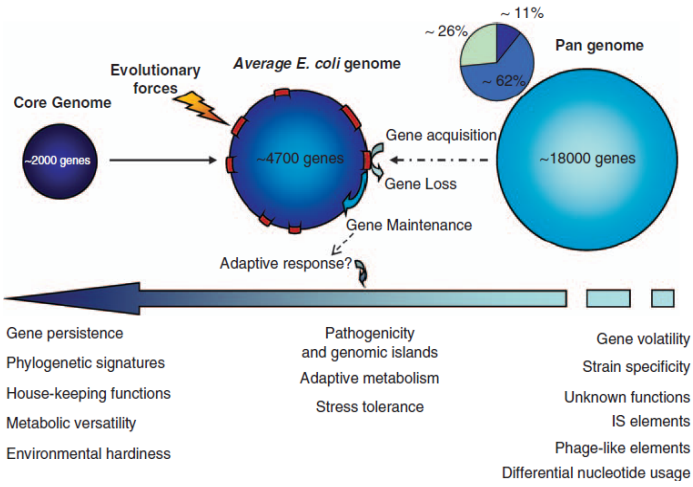


Fig. 7.2: A multitude of evolutionary forces shape the E. coli genome. E. coli strains possess a core of about 2,000 genes which equip them with a versatile metabolism. The pangenome is constituted of about 18,000 genes, of which 11% belong to the core (dark blue), 62% is so-called of persistent genes (blue), 26% can be considered "accessory" genes (pale blue). The arrow represents a hypothetical gradient in which E. coli genome features that are most likely associated with volatile, persistent and core genes are shown (van Elsas et al., 2011).

7.1. Identification and characterisation of genetic determinants encoding virulence factors: the example of SubAB in the LEE-negative VTEC strains

VTEC isolates lacking the LEE locus have been raising interest in the scientific community since they are able to induce disease in human in spite they are not able to cause the typical A/E lesion to the gut mucosa. Multilocus sequence typing (MLST) based on housekeeping genes has shown that LEE-negative VTEC isolates are evolutionary divergent (Tarr *et al.*, 2008). A more recent study based on the whole genome sequencing (WGS) of nine LEE-negative VTEC isolates confirmed their phylogenetic diversity and highlighted the lack of genes in common between such isolates (Steyert *et al.*, 2012).
One of the virulence factor recently described in LEE-negative VTEC is the SubAB. This is a AB_5 toxin showing a Subtilase-like serine protease activity. First described in the large virulence plasmid of a VTEC O113:H21, the SubAB was later identified in other LEE-negative VTEC strains. It has been hypothesised that this toxin may contribute to the VTEC pathogenicity by a synergic action with VTs. SubAB cause a cytopathic effect on Vero cells and induces, if injected intraperitoneally in mice, the typical triad of HUS symptoms. More recently, *subAB* genes have been described in two *E. coli* strains negative for the production of VTs (Tozzoli *et al.*, 2010). In the latter strains, the Subtilase was encoded by an allelic variant

of such genes, termed *subAB*$_2$, which share 90% homology with the corresponding variant in the prototype strain and is carried by an 8 kb-long PAI (SE-PAI) located on the chromosome, which also hosts the *tia* gene, encoding an invasion determinant, and the *shiA* gene, encoding an immunomodulator in *S. flexnerii*. This PAI was shown to be integrated in the PheV locus, a hot-spot integration site for many PAIs in *E. coli*, including the LEE locus. It could be hypothesised that VTEC lacking the LEE locus and carrying the SE-PAI may represent the result of a mutual competition between the two islands for the same integration site, thus explaining the strong association of SE-PAI with LEE-negative strains.
At least two *subAB* variants have been individuated so far (Paton *et al.*, 2001, Michelacci *et al.*, 2013). Recently, an additional subtype of *subAB,* termed *subAb*$_{2\text{-}2}$ and displaying more than 99.0 % homology to *subAB*$_2$ has been described. It is conveyed by a chromosomally located PAI different from the SE-PAI and termed OEP-locus, located in the proximity of an operon encoding a type I secretion system (Fig. 7.3) (Funk *et al.*, 2013).
The possible reservoir of *subAB*-positive LEE-negative VTEC strains has not been investigated in detail. Some authors reported the association of *E. coli* strains carrying these genes with small ruminants, such as deer and sheep (Sánchez *et al.*, 2012). A PCR-based tool able to discriminate between the prototype *subAB* and the *subAB*$_2$ variant was used to investigate their distribution in a panel of *E. coli* strains isolated from human and animal sources. *subAB* genes were confirmed to be only present in LEE-negative VTEC strains. Additionally, their presence was confirmed in 70% of LEE-negative VTEC isolated from human diarrhoea assayed in the study, thus confirming that SubAB may indeed play a role in the disease caused by these VTEC strains (Michelacci *et al.*, 2013).
Interestingly, the vast majority of the *subAB*-positive strains carried the *subAB*$_2$ variant and seemed to possess the entire SE-PAI. The presence of this PAI was also observed in 86% of the animal LEE-negative VTEC strains investigated.
Therefore, small ruminants appear to be an important animal reservoir for the *subAB*$_2$-positive LEE-negative VTEC strains, implying a probable zoonotic origin also for this VTEC group as it has been proposed for the typical EHEC and ascertained for the VTEC O157:H7.
The SE-PAI could be involved in the pathogenetic mechanism in humans, extending the concept of VTEC "virulome" to this MGE. Even though the function of the other virulence genes carried by this PAI, such as the *tia* and *shiA* genes, has to be determined, their products may well be involved in the pathogenetic

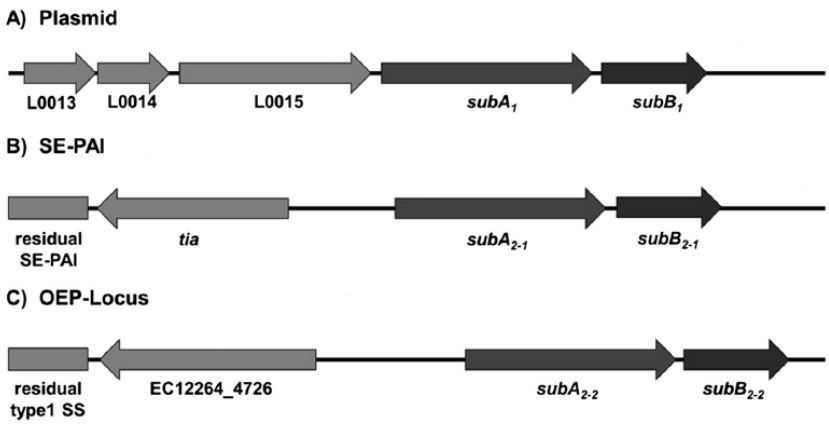


Fig. 7.3: Schematic illustration of the different loci of subAB genes. A) Plasmidic locus of the subAB$_1$ *variant of E. coli O113:H21 strain; B) SE-PAI of the E. coli O78:H- strain carrying the subAB*$_2$ *variant; C) OEP-locus of the subAB*$_{2\text{-}2}$ *variant in the O76:H- strain in which it was described (adapted from Funk et al., 2013).*

mechanism. The *shiA* gene has a counterpart in *S. flexnerii* where its product acts as an immunomodulator, while the product of an homologue of *tia* gene in ETEC has been demonstrated to be involved in the invasion of the host cells upon infection (Fleckenstein *et al.*, 1996). It could be hypothesised that the proteins encoded by these two genes may have a role in the colonization process operated by the SE-PAI positive VTEC strains, providing an alternative to the attaching and effacing lesion caused by EHEC.

7.2. Identification and characterisation of genetic determinants encoding virulence factors: the example of *toxB* in LEE-positive VTEC strains

VTEC strains possessing the LEE locus have been historically considered as the most pathogenic ones. The current VTEC classification proposed by Karmali and colleagues (Karmali et al., 2003) is based on phenotypic and molecular characteristics of the strains and on the clinical features of the associated disease. According to this model, VTEC strains are categorised into seropathotypes (SPTs) indicated with letters from A to E, with a decreasing rank of pathogenicity. VTEC O157 constitute the SPT A, while VTEC strains belonging to O26, O111, O103, O145 and O121, positive for the presence of the LEE locus and frequently associated with outbreaks, constitute the SPT B. SPT C and D include LEE-negative VTEC strains rarely associated with HUS or causing mild diarrhoea, respectively. Finally, SPT E strains predominantly circulate in the animal reservoir and have not been reported so far as cause of disease in the human host (Karmali *et al.*, 2003).

ToxB is reported as being one of the factors composing the most pathogenic VTEC (SPT A and B) virulence repertoire. In fact, it has been shown to be strictly present in LEE-positive VTEC strains (Tozzoli *et al.*, 2005, Bosilevac *et al.*, 2011, Horcajo *et al.*, 2012). *toxB* is a 9.5 Kb long gene whose product share 28% identity and 47% similarity in the aminoacidic sequence with the protein encoded by *efa1/lifA* gene (Tozzoli *et al.*, 2005). The presence of *toxB* gene is increasingly considered a marker of the large virulence plasmid of EHEC, and it has been ascertained in the large virulence plasmid of VTEC O157, VTEC O26, in a VTEC O145:NM strain (Venturini *et al.*, 2010, Fratamico *et al.*, 2011) and more recently in the large plasmid of a VTEC O111 strain from a human case of diarrhoea (Michelacci *et al.*, 2014). As with many other virulence genes of VTEC, *toxB* exists in at least two allelic variants displaying a distribution non-randomly segregating with the serogroup. In particular, the $toxB_1$ allele seems to be present exclusively in VTEC O157, while the $toxB_2$ has been identified in the VTEC strains belonging to SPT B, although this gene may be absent in some VTEC strains belonging to SPT B and still carrying the large virulence plasmid. The latter observation could be explained by the ability of this gene to self-mobilize, being surrounded by IS3-like elements (Tozzoli *et al.*, 2005).

The strict association of *toxB* gene with VTEC strains carrying the LEE locus strengthens the hypothesis proposed by some authors of its role in enhancing the effect of the A/E lesion, particularly by influencing the initial stages of adherence (Tatsuno *et al.*, 2000). Accordingly, VTEC O157 strains in which *toxB* was mutated, produced a reduced number of micro-colonies on HeLa cultured cells, but were still able to nucleate filamentous actin under the sites of attachment, suggesting that the LEE locus-mediated rearrangement of cytoskeletal actin can still occur in the absence of a functional *toxB* (Fig. 7.4) (Stevens *et al.*, 2004).

In addition to the observed reduction of the micro-colonies, the removal of the entire pO157 plasmid from a VTEC O157 caused reduced levels of EspA, EspB, and Tir proteins. The level of these T3SS-proteins was restored in the presence of a plasmid containing ToxB but not in the presence of plasmids carrying other pO157-borne genes, suggesting a probable influence of the ToxB on the activity of the T3SS (Tatsuno *et al.*, 2000).

A function in the adhesion of *E. coli* strains causing the attaching and effacing lesion has also been hypothesised for the product of the *efa1/lifA* gene, a gene harboured on the PAI OI-122, which has been

shown to contribute to the adherence of EPEC strains on cultured cells (Nicholls *et al.*, 2000). It has been proposed that the two proteins, Efa1/lifA and ToxB, could have a synergistic action in the influence on the T3SS (Stevens *et al.*, 2004).

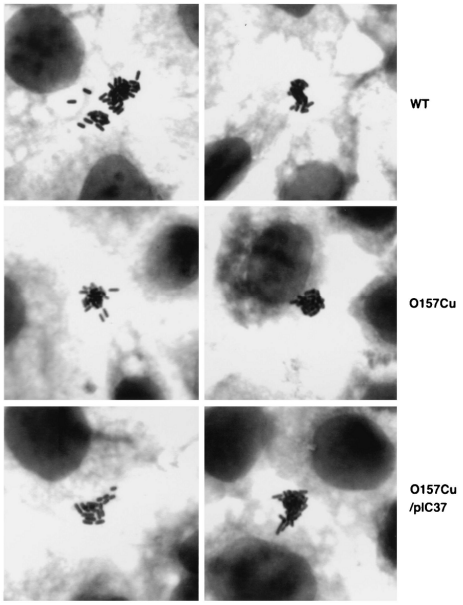


Fig. 7.4: Adherence phenotypes of a pO157-cured mutant (O157Cu) and O157Cu/pIC37 (toxB gene) (Tatsuno et al., 2000).

It has to be noted that similarly to how observed for *toxB* also *efa1/lifA* gene sequence presents differences in VTEC strains belonging to SPT A or SPT B (Morabito *et al.*, 2003, Tozzoli *et al.*, 2005). Efa1/lifA is full length in the SPT B strains while it is truncated in VTEC O157 strains where only the 3 primus extremity is present. Nevertheless, it seems that this portion of the gene still retains at least part of its functionality (Stevens *et al.*, 2004). It could be hypothesised that the different *toxB* allelic variants could have evolved to complement different activity levels of the two forms of the Efa1/lifA.

7.3 VT-phages from VTEC isolated from severe human disease and their role in the host biology

VT-bacteriophages are responsible for the dissemination of VT-coding genes (*vtx).* VTs are AB_5 toxins blocking the protein synthesis and causing cell death in target organs. This can lead to the development of the most severe forms of VTEC infections, haemorrhagic colitis (HC) and haemolytic uremic syndrome (HUS).

As part of my PhD project, I focused my attention on the VT-phages present in VTEC O157:H7 strains associated with human disease and on those present in the recently described *E. coli* pathotype Enteroaggregative Haemorrhagic *E. coli* (EAHEC), with the aim to investigate the bases of the emergence of the different VTEC groups.

We described that the region comprised between the genes *gam* and *cII* was a marker allowing the specific identification of the VT2-phage of VTEC O157 significantly associated with human disease compared to those present in the animal reservoir. Such a VT2-phage was termed Φ-8 (Chapter 5, publication 1).

In spite of the initial identification of Φ-8 in VTEC O157 strains belonging to phage type (PT) 8, we could demonstrate that this type of phage was significantly associated to strains isolated from the human disease regardless their PT and that it was also present in VTEC O157 of PT 21/28 predominantly isolated in the United Kingdom from human disease and from cattle presenting high level excretion of such VTEC strains in their faeces, the so-called super-shedders cattle.

We have described that, at least in the laboratory conditions, Φ-8 phage does not induce higher levels of VT2 mRNA compared to other types of VT-phage. At the same time we found that its presence may influence the ability of VTEC O157 strains to cause the A/E lesion, since it seems to induce down-regulation of the main promoter of the LEE locus in a recombinant *E. coli* system. Accordingly, we have observed that VTEC O157 strains possessing Φ-8 naturally showed reduced levels of the secreted EspD protein, a T3SS marker, when compared to VTEC O157 carrying a BP933W-like phage. Similar results were obtained with Φ-8 positive VTEC O157 strains belonging to PT 21/28 and isolated from super-shedders animals, thus strengthening the hypothesis that Φ-8 phage may be involved in the VTEC colonization mechanism of the host by operating a fine regulation of some of the LEE genes. The finding of a high prevalence of Φ-8 in VTEC O157 strains from super-shedder animals is interesting and might explain to a certain extent the comparably higher prevalence of this VT2-phage in VTEC O157 from human disease. As a matter of fact, the presence of Φ-8 might favour the prolonged colonization of the animal reservoir by diminishing the impact of the A/E lesion on the gut, eventually leading to super-shedding status, in turn establishing amplification cycles in the farm and resulting in an increased exposure of the human host to Φ-8-positive VTEC O157.

Among other possible factors contributing to the fine-tuned regulation of the colonization of the animal reservoir by VTEC O157, we considered the *rpoS* gene. RpoS is an alternative sigma factor intervening in the stationary-phase stress response by adjusting the regulation of a wide range of genes involved in the acid-resistance, osmotic stress, heat-shock and starvation (van Hoek *et al.*, 2013). It has been shown that *E. coli* strains with attenuated RpoS levels have lower levels of resistance to external stresses but have some increased nutritional abilities (van Hoek *et al.*, 2013). Additionally, some studies reported that RpoS negatively regulates the expression of the LEE locus genes, with VTEC O157 strains mutated in *rpoS* showing a higher expression of virulence genes related to the colonization mechanism (Dong *et al.*, 2010). The investigation of *rpoS* sequence in a panel of 40 VTEC O157 strains carrying Φ-8 or a BP933W-like VT-phage showed that most of the components of the former group clustered apart from the latter with few exceptions, when clustalW analysis was performed (unpublished data; Fig. 7.5).

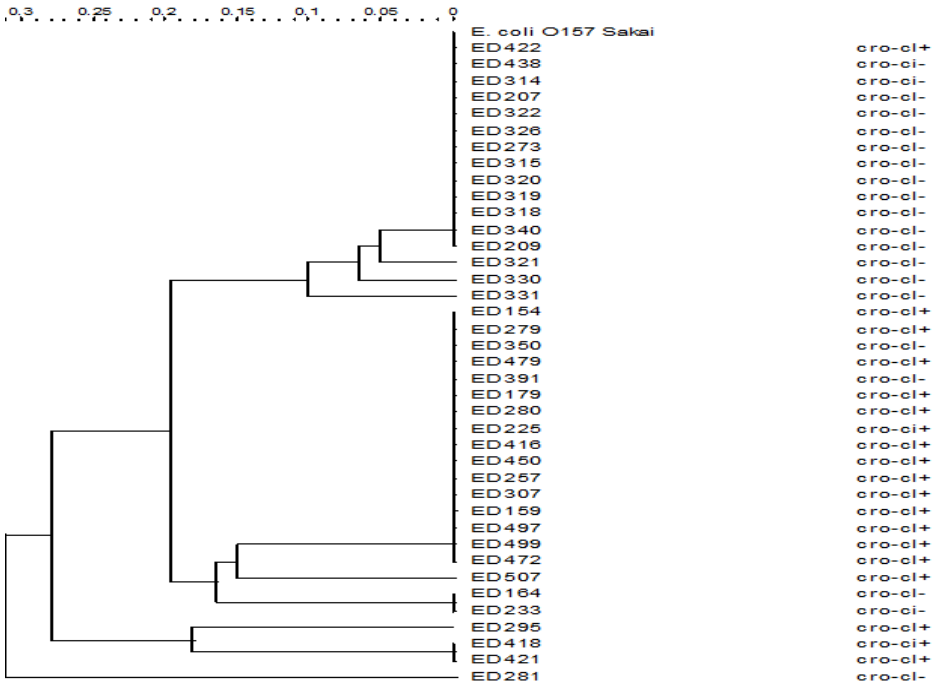


Fig. 7.5. rpoS gene sequences clustering. ClustalW analysis of rpoS gene sequences from VTEC O157 strains carrying a Φ-8 phage (cro-cI+) or a BP933W-like VT phage (cro-cI-).

With the aim to understand the basis of the association between the VT-phage carried and *rpoS* allele, we assessed the expression of *rpoS* in the two groups.
We observed that the level of mRNA of the *rpoS* alleles did not differ significantly in the two groups, suggesting that the reason for such an association could be in the protein function or somewhere else (unpublished data).
These observation, although not conclusive, strengthen the proposed association between Φ-8 and ecological factors favouring the success of VTEC O157 in establishing in specific niches, e.g. the animal reservoir and the farm environment.
The proposed association of VTEC O157 Φ-8-positive strains with a more efficient animal colonization is also strengthened by the recent observation that these strains show a significantly higher adhesion level on a calf intestinal epithelial cell line (CIEB) ($p<0.001$) compared to VTEC O157 carrying a VT-phage similar to the VTEC O157 EDL933 strain, while the two groups did not show differences in the adhesion levels when the assay was carried using the human epithelial colorectal adenocarcinoma Caco-2 cell line (unpublished data). By using the same experimental model, we also assessed that a VTEC O157 strain from which the Φ-8 phage was experimentally removed, was impaired in adhering to the CIEB cells with respect to its parental strain.
Altogether, these findings support the hypothesis for a role of VT-phages not only limited to the carriage and spreading of *vtx*-genes, but also in influencing the regulation of other virulence factors and more in general to manipulate the host biology conferring to VTEC O157 a better fitness in the reservoir and the environment.
We also investigated the relationship between the VT-phages and their bacterial host. The recently described Enteroaggregative Haemorrhagic *E. coli* (EAHEC) have been proposed to arise from Enteroaggregative *E. coli* (EAggEC) strains following the acquisition of a VT2-phage.
We have sequenced and annotated the genome of the VT2-phage (Phi-191) carried by the first EAHEC strain ever described, isolated in the '90s during a small outbreak of HUS in France (Morabito *et al.*, 1998) and found that it was more than 99% homologous to that of the VT2-phage carried by the EAHEC O104:H4 that have caused the large German outbreak in 2011, isolated 20 years later (Grande *et al.*, 2014). This is noteworthy, since VT-phages are known to be hugely variable (Muniesa *et al.*, 2014). At the same time, the WGS of Phi-191 seemed to be largely divergent from the VT2-phage present in the EAHEC O111:H21 isolated from a HUS case occurred in Northern Ireland in 2012. If the first finding is may be in favour of a strong selective pressure acting on the VT-phage limiting their indiscriminate access to EAggEC, the identification of a VT2-phage completely different in the EAHEC O111:H21 indicates that at least two types of VT-phages have been able to stably infect EAggEC recipients and suggesting that such types of phages may be combined into a somehow homogeneous group. In order to ascertain if it was the case we have aligned the sequence of the Phi-191 with those of VT-phages from other VTEC available in GenBank, including EAHEC or typical VTEC, and could observe that the VT-phages of EAHEC had in common the presence of two sequence blocks that were not present in the other VT-phages. One of those DNA stretches corresponded to a gene apparently coding for a phage tail fiber, which could also be detected in the short reads composing the genome of the O111:H21 EAHEC isolated in Northern Ireland in 2012 (Grande *et al.*, 2014) and in the whole genome sequence of a EAHEC O127:H4 strain isolated in 2013 in Northern Italy (unpublished data). Since it has been hypothesised that the infection with a lambdoid phage can be mediated by the cross-talking between the bacterium and the phage and tail fibers are part of this selection process (Werts *et al.*, 1994), we hypothesised that such a fiber might be part of the VT-phage host specificity machinery.
The observed tail fiber-encoding gene has a peculiar hit when searched across the GenBank, which corresponds to a tail fiber from a phage isolated from *Shigella boydii*. This finding is interesting and could

lead to the hypothesis that EAHEC and typical VTEC might have originated from multiple acquisition events of VT-phages. In the typical VTEC, they could have undergone a highly divergent evolution originating multiple types of phages with different tail fibers given the wider circulation of these strains into animal reservoirs. Alternatively, EAHEC might be restricted to one single host, explaining the high stability observed in their VT-phage genomes. Such a scenario is also compatible with the inter-human circulation of EAggEC, which also seems to explain the different epidemiology observed for the infection caused by EAHEC and typical VTEC. In fact, the inter-human circulation of EAHEC may justify the occurrence of sporadic cases and small outbreaks of disease, as we have observed from the 1992 to 2011 and up to nowadays. The German outbreak of EAHEC O104:H4 occurred in 2011 represents an exception to this scheme, however it could be explained by the accidental contamination of the fenugreek seeds used for sprouting, reasonably from a human worker, in one of the many segments of the seeds production chain. The use of such contaminated seeds and the concomitant presence of many people gathered together for the Hamburg harbour celebration that have consumed the sprouts, could explain the unusual high number of cases of infection occurred during the German outbreak.
The original source of the VT-phages, though, still has to be identified.

7.4 Emerging pathogenic *E. coli* threats: who's the next?

The large 2011 German outbreak had a deep impact on the EU and forced the scientific community to start re-thinking the existing relationships between the different groups of pathogenic *E. coli*. One of the largest and undoubtedly the most severe *E. coli* outbreak ever experienced by the humankind was indeed caused by an *E. coli* strain presenting an unusual combination of virulence factors overlapping two different groups: VTEC and Enteroaggregative *E. coli*. As a matter of fact, the causative agent of this outbreak presented the backbone genetic organization and phenotypic features of an EAggEC strain which acquired the ability to produce the Verocytotoxin following a lysogenization event with a VT-phage (Bielaszewska *et al.*, 2011). Such an unusual combination of virulence features was not previously unseen, though. Some sporadic cases and a small HUS outbreak caused by EAHEC strains could be identified either retrospectively *via* the scientific literature (Morabito *et al.*, 1998, Iyoda *et al.*, 2000, Sca*via et al.*, 2011) or actively observed from the German outbreak on, such as the EAHEC case of infection described in 2012 in the UK (Dallman *et al.*, 2012), or the small outbreak of HUS occurred in Italy in 2013 (Tozzoli *et al.*, 2014a) . The identification of EAHEC infections occurred in a 20 years' time span seems to indicate that EAHEC may be not just chimeric strains deriving from accidental events of VT-phages acquisition but rather a new pathogenic group of *E. coli* in its own right.
An event of VT-phage acquisition is in principle possible for any *E. coli* strain. This possibility has been assessed either experimentally or by direct observation of the existence of bacterial strains isolated from cases of disease in humans displaying a number of different genetic backgrounds. In the laboratory conditions, pathogenic *E. coli* strains belonging to different pathotypes including EPEC, VTEC, EAggEC, ETEC and EIEC as well as ExPEC and *E. coli* strains isolated from healthy individuals have been shown to successfully acquire and in some cases to stably maintain a VT-phage (Schmidt *et al.*, 1999, Beutin *et al.*, 2013, Tozzoli *et al.*, 2014a). The laboratory observations are corroborated by a number of reports in the literature describing the presence of VT-phage or the toxin production in various Enterobacteriaceae isolated from cases of human infections including at least three different pathotypes of *E. coli,* ExPEC, EAggEC and ETEC (Karch *et al.*, 1999, Tozzoli *et al.*, 2014a), *Shigella dysenteriae type I*, *Shigella sonnei, Shigella flexnerii* (Strauch *et al.*, 2001, Gray *et al.*, 2014), *Citrobacter freundii* (Schmidt *et al.*, 1993, Tschape *et al.*, 1995) and *Enterobacter cloacae* (Paton *et al.*, 1996).
Taken together, all these observations suggest that there is a pool of VT-phages that are continuously acquired and lost from bacterial hosts. These phages may sometimes be stably maintained in a lysogenic state in the host chromosome originating a clone of VT-producing bacteria. When it happens to be an *E.*

coli strain, it can be a harmless VTEC or, depending on the availability of a genetic background associated with the capacity to cause disease, namely the presence of a colonization machinery, it may give origin to a strain with augmented virulence. It can be assumed that such a scenario caused the EHEC and EAHEC groups to emerge from EPEC and EAggEC, respectively, as established pathogenic groups of *E. coli*.

7.5 Concluding Remarks

Diarrheagenic *E. coli* (DEC), including Verocytotoxin-producing *E. coli* (VTEC), are a significant public health issue worldwide. The management of the infections caused by these bacterial pathogens is complicated by their extreme heterogeneity, including strains causing a plethora of symptoms spanning from uncomplicated diarrhoea to life-threatening systemic sequelae such as the HUS. This picture is further intricate because of the complex epidemiology of these pathogens. Some of the DEC groups have an inter-human circulation and an oral-faecal route of transmission of the infections, such as the Enteroaggregative *E. coli* (EAggEC), typical Enteropathogenic *E. coli* (tEPEC) and Enteroinvasive *E. coli* (EIEC), and are of great concern especially in developing countries. Enterotoxigenic *E. coli* (ETEC), atypical Enteropathogenic *E. coli* (aEPEC) and Verocytotoxin-producing *E. coli* (VTEC) are mainly food-borne transmitted, and may have human or animal reservoirs. It has to be noted that a zoonotic connection has been exclusively recognized for the latter group, although it has been proposed that also aEPEC infections may be acquired through contacts involving animals or food of animal origins. aEPEC and VTEC infections are common in industrialized countries where their diffusion benefits from the large distribution of foodstuffs that accounts for the occurrence of outbreaks sometimes involving a high number of cases. In these geographic areas these infections, especially those caused by VTEC, represent a major public health issue, due to the severity of the disease caused.

Finally, the recently recognized pathotype of Enteroaggregative Haemorrhagic *E. coli* (EAHEC) highlighted that also the infections diffused by those DEC groups with an inter-human circulation may actually be placed at the human-food interface. As a matter of fact, these pathogenic *E. coli* may have the chance to enter the food chain *via* the contamination operated by the food handlers or, as it was the case of EAHEC O104:H4 that caused the German outbreak in 2011, during one of the many phases characterising the production of the seeds used to produce sprouts.

The results presented in this piece of research attempted to address some of the open questions in the field of pathogenic *E. coli*, especially those regarding VTEC as a pathotype. The *leitmotif* connecting all the publications is that the pathogenic mechanism of all VT-producing *E. coli* causing disease to humans is based on two aspects: the VT-production and the colonization mechanism.

The most severe forms of infections, HC and HUS, are caused by the action of VTs on the microvasculature of the gut mucosa and renal glomerular cells. For the action of VTs to occur, however, the VTEC must stay in the intestine long enough to allow the transfer of the toxin from the gut lumen to the blood vessels. This is achieved through an efficient colonization. Otherwise, the bacteria are cleared with the faeces and the VTs will not reach their final targets. The papers presented in this piece of research dealt with pathogenic *E. coli* strains presenting different rank of pathogenicity, spanning from the LEE-negative VTEC to VTEC possessing the LEE locus and causing HUS up to the most virulent *E. coli* ever described, EAHEC, and highlighted that the virulence potential of the VTs is expressed in parallel with a the ability to colonize the host gut (Fig. 7.6).

A number of evidences have been accumulated during my PhD thesis that strengthen the inter-relationships between the presence of a VT-phage by *E. coli* and the ability to colonize the host intestine including:

- A new VTEC pathogenicity island, the SE-PAI, has been described as part of the “virulome” of VTEC lacking the LEE locus (SPT C) and included determinants encoding both a VT-related toxin, the Subtilase, and proteins similar to colonization factors described in other *E. coli* groups

(Tia) or other enteric pathogens (ShiA). Such factors may contribute to the colonization mechanism as an alternative to the A/E lesion in LEE-negative VTEC strains causing human disease.

- The study on *toxB* allelic variants indicated that this gene is one of the factors composing the "accessory" virulence repertoire in most of the pathogenic VTEC strains belonging to SPT A and B. The strict association of ToxB with VTEC LEE positive strains supports its proposed role in enhancing the effect of the A/E lesion and it could be hypothesised that different ToxB variants could have different activities in contributing to the pathogenic potential of strains belonging to SPT A and SPT B.

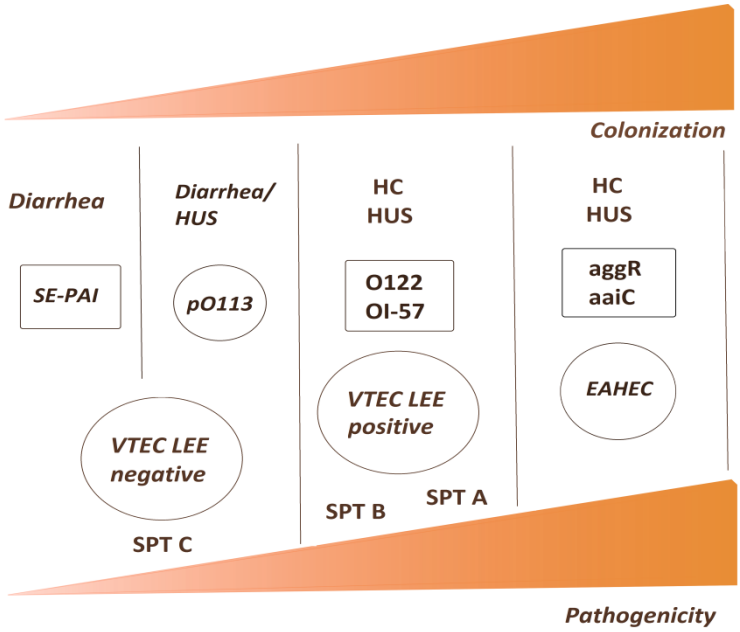


Fig. 7.6. Severity of the symptoms associated with the infection with different VTEC groups alongside with their capability to colonize the host intestinal mucosa.

- Beside the colonization, also the type of mobile genetic elements conveying the VT-coding genes may confer more virulence or a higher chance to cause disease in human beings. The VT-phage Φ-8 is in fact associated to VTEC O157 strains isolated from severe human infections. Such an effect may in turn be related with the colonization, as we could demonstrate that this phage has an influence on the T3SS activity through the down-regulation of the LEE locus genes. In this particular case, though, the influence may be on the colonization of the animal reservoir, given the association of Φ-8 with VTEC O157 strains from super-shedders animals.

Finally we have obtained evidences suggesting that the VT-phages may have evolved or even still be evolving with the bacterial host. In fact we have observed that:

- The whole genome sequence of VT-phage (Phi-191) from the first EAHEC strain ever described was more than 99% homologous to that of the VT-phage isolated from the EAHEC strain that have caused the German outbreak 20 years later. This observation is puzzling being the phage described as highly variable organisms. Moreover, we observed that the VT-phage present in another EAHEC, the EAHEC O111:H21 isolated in Northern Ireland in 2012, was completely different, suggesting that at least two VT-phages intervened in the emergence of this new VTEC group.

- VT-phages infecting EAggEC seems to be characterised by the presence of a peculiar tail fiber gene that could be involved in the host-pathogen recognition process and in the definition of the host spectrum specificity.
- Different types of VT-phages seem to be able to infect a heterogeneous group of pathogenic and non-pathogenic *E. coli*, belonging to both DEC and ExPEC groups as well as other Enterobacteriaceae, showing a host range broader than expected. This raises the hypothesis that probably any Enterobacteria equipped with a colonization apparatus could be potentially infected by a VT-phage generating clones with augmented virulence potential for humans, as it happened with the VT-producing *Enterobacter cloacae* isolated from an HUS patient or the EAHEC that caused the 2011 German outbreak.

REFERENCES

Ahmed, S.A., Awosika, J., Baldwin, C., Bishop-Lilly, K.A., Biswas, B., Broomall, S.*, et al.* (2012). Genomic comparison of Escherichia coli O104:H4 isolates from 2009 and 2011 reveals plasmid, and prophage heterogeneity, including shiga toxin encoding phage stx2. *PLoS One* **7,** e48228.

Andreoli, S.P., Trachtman, H., Acheson, D.W., Siegler, R.L. and Obrig, T.G. (2002). Hemolytic uremic syndrome: epidemiology, pathophysiology, and therapy. *Pediatr Nephrol* **17,** 293-298.

Beutin, L., Hammerl, J.A., Reetz, J. and Strauch, E. (2013). Shiga toxin-producing Escherichia coli strains from cattle as a source of the Stx2a bacteriophages present in enteroaggregative Escherichia coli O104:H4 strains. *Int J Med Microbiol* **303,** 595-602.

Bhan, M.K., Khoshoo, V., Sommerfelt, H., Raj, P., Sazawal, S. and Srivastava, R. (1989). Enteroaggregative Escherichia coli and Salmonella associated with nondysenteric persistent diarrhea. *Pediatr Infect Dis J* **8,** 499-502.

Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Fruth, A., Bauwens, A.*, et al.* (2011). Characterisation of the Escherichia coli strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. *Lancet Infect Dis* **11,** 671-676.

Boisen, N., Nataro, J.P. and Krogfelt, K.A. (2014) Enteroaggregative Escherichia coli and Disease, in Pathogenic Escherichia coli, Cellular and Molecular Microbiology.

Bosilevac, J.M. and Koohmaraie, M. (2011). Prevalence and characterisation of non-O157 shiga toxin-producing Escherichia coli isolates from commercial ground beef in the United States. *Appl Environ Microbiol* **77,** 2103-2112.

Bower, J.R., Congeni, B.L., Cleary, T.G., Stone, R.T., Wanger, A., Murray, B.E.*, et al.* (1989). Escherichia coli O114:nonmotile as a pathogen in an outbreak of severe diarrhea associated with a day care center. *J Infect Dis* **160,** 243-247.

Boyd, E.F. and Brüssow, H. (2002). Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. *Trends Microbiol* **10,** 521-529.

Brunder, W., Karch, H. and Schmidt, H. (2006). Complete sequence of the large virulence plasmid pSFO157 of the sorbitol-fermenting enterohemorrhagic Escherichia coli O157:H- strain 3072/96. *Int J Med Microbiol* **296,** 467-474.

Brzuszkiewicz, E., Gottschalk, G., Ron, E., Hacker, J. and Dobrindt, U. (2009). Adaptation of Pathogenic E. coli to Various Niches: Genome Flexibility is the Key. *Genome Dyn* **6,** 110-125.

Cambray, G., Guerout, A.M. and Mazel, D. (2010). Integrons. *Annu Rev Genet* **44,** 141-166.

Campbell, A. (1996) Bacteriophages, In: Escherichia coli and Salmonella: Cellular and Molecular Biology., ASM Press, Washington, DC.

Caprioli, A., Morabito, S., Brugère, H. and Oswald, E. (2005). Enterohaemorrhagic Escherichia coli: emerging issues on virulence and modes of transmission. *Vet Res* **36,** 289-311.

Clarke, S.C. (2001). Diarrhoeagenic Escherichia coli--an emerging problem? *Diagn Microbiol Infect Dis* **41,** 93-98.

Coombes, B.K., Gilmour, M.W. and Goodman, C.D. (2011). The evolution of virulence in non-o157 shiga toxin-producing Escherichia coli. *Front Microbiol* **2,** 90.

Dallman, T., Smith, G.P., O'Brien, B., Chattaway, M.A., Finlay, D., Grant, K.A. and Jenkins, C. (2012). Characterisation of a verocytotoxin-producing enteroaggregative Escherichia coli serogroup O111:H21 strain associated with a household outbreak in Northern Ireland. *J Clin Microbiol* **50,** 4116-4119.

Del Canto, F. and Torres, A.G. (2014) "Colonization Factor Antigens of Enterotoxig*enic Escherichia* coli" in P*athogenic Escher*ichia coli, Molecular and Cellular Microbiology. Caister Academic Press, pp. 201-216.

Dobrindt, U., Chowdary, M.G., Krumbholz, G. and 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. and Hacker, J. (2004). Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* **2,** 414-424.

Dong, T. and Schellhorn, H.E. (2010). Role of RpoS in virulence of pathogens. *Infect Immun* **78,** 887-897.

Donnenberg, M.S. (1995) Enteropathogenic Escherichia coli infections of the gastrointestinal tract., Raven Press, Ltd, New York.

Donnenberg, M.S. and Whittam, T.S. (2001). Pathogenesis and evolution of virulence in enteropathogenic and enterohemorrhagic Escherichia coli. *J Clin Invest* **107,** 539-548.

Dudley, E.G., Thomson, N.R., Parkhill, J., Morin, N.P. and Nataro, J.P. (2006). Proteomic and microarray characterisation of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative Escherichia coli. *Mol Microbiol* **61,** 1267-1282.

Dutta, S., Guin, S., Ghosh, S., Pazhani, G.P., Rajendran, K., Bhattacharya, M.K., *et al.* (2013). Trends in the prevalence of diarrheagenic Escherichia coli among hospitalized diarrheal patients in Kolkata, India. *PLoS One* **8,** e56068.

ECDC (2014) Annual Epidemiological Report 2013. Reporting on 2011 surveillance data and 2012 epidemic intelligence data. ECDC.

Fleckenstein, J.M., Kopecko, D.J., Warren, R.L. and Elsinghorst, E.A. (1996). Molecular characterisation of the tia invasion locus from enterotoxigenic Escherichia coli. *Infect Immun* **64,** 2256-2265.

Foxman, B. (2010). The epidemiology of urinary tract infection. *Nat Rev Urol* **7,** 653-660.

Frank, C., Werber, D., Cramer, J.P., Askar, M., Faber, M., an der Heiden, M., *et al.* (2011). Epidemic profile of Shiga-toxin-producing Escherichia coli O104:H4 outbreak in Germany. *N Engl J Med* **365,** 1771-1780.

Franz, E., Delaquis, P., Morabito, S., Beutin, L., Gobius, K., Rasko, D.A., *et al.* (2014). Exploiting the explosion of information associated with whole genome sequencing to tackle Shiga toxin-producing Escherichia coli (STEC) in global food production systems. *Int J Food Microbiol* **187,** 57-72.

Fratamico, P.M., Yan, X., Caprioli, A., Esposito, G., Needleman, D.S., Pepe, T., *et al.* (2011). The complete DNA sequence and analysis of the virulence plasmid and of five additional plasmids carried by Shiga toxin-producing Escherichia coli O26:H11 strain H30. *Int J Med Microbiol* **301,** 192-203.

Funk, J., Stoeber, H., Hauser, E. and Schmidt, H. (2013). Molecular analysis of subtilase cytotoxin genes of food-borne Shiga toxin-producing Escherichia coli reveals a new allelic subAB variant. *BMC Microbiol* **13,** 230.

Gaastra, W. and Svennerholm, A.M. (1996). Colonization factors of human enterotoxigenic Escherichia coli (ETEC). *Trends Microbiol* **4,** 444-452.

Grande, L., Michelacci, V., Tozzoli, R., Ranieri, P., Maugliani, A., Caprioli, A. and Morabito, S. (2014). Whole genome sequence comparison of vtx2-converting phages from Enteroaggregative Haemorrhagic Escherichia coli strains. *BMC Genomics* **15,** 574.

Gray, M.D., Lampel, K.A., Strockbine, N.A., Fernandez, R.E., Melton-Celsa, A.R. and Maurelli, A.T. (2014). Clinical Isolates of Shiga Toxin 1a-Producing Shigella flexneri with an Epidemiological Link to Recent Travel to Hispañiola. *Emerg Infect Dis* **20,** 1669-1677.

Griffin, P.M. and Tauxe, R.V. (1991). The epidemiology of infections caused by Escherichia coli O157:H7, other enterohemorrhagic E. coli, and the associated hemolytic uremic syndrome. *Epidemiol Rev* **13,** 60-98.

Hacker, J., Blum-Oehler, G., Mühldorfer, I. and Tschäpe, H. (1997). Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol Microbiol* **23,** 1089-1097.

Hall, R.M. (2012a). Integrons and gene cassettes: hotspots of diversity in bacterial genomes. *Ann N Y Acad Sci* **1267,** 71-78.

Hall, R.M. (2012b) Integrons and gene cassettes: hotspots of diversity in bacterial genomes., Ann. N.Y. Acad. Sci.

Harrington, S.M., Dudley, E.G. and Nataro, J.P. (2006). Pathogenesis of enteroaggregative Escherichia coli infection. *FEMS Microbiol Lett* **254,** 12-18.

Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., *et al.* (2001). Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res* **8,** 11-22.

Hazen, T.H., Sahl, J.W., Redman, J.C., Morris, C.R., Daugherty, S.C., Chibucos, M.C., *et al.* (2012). Draft genome sequences of the diarrheagenic Escherichia coli collection. *J Bacteriol* **194,** 3026-3027.

Hebbelstrup Jensen, B., Olsen, K.E., Struve, C., Krogfelt, K.A. and Petersen, A.M. (2014). Epidemiology and clinical manifestations of enteroaggregative Escherichia coli. *Clin Microbiol Rev* **27,** 614-630.

Heuvelink, A.E., van Heerwaarden, C., Zwartkruis-Nahuis, J.T., van Oosterom, R., Edink, K., van Duynhoven, Y.T. and de Boer, E. (2002). Escherichia coli O157 infection associated with a petting zoo. *Epidemiol Infect* **129,** 295-302.

Ho, T.D., Davis, B.M., Ritchie, J.M. and Waldor, M.K. (2008). Type 2 secretion promotes enterohemorrhagic Escherichia coli adherence and intestinal colonization. *Infect Immun* **76,** 1858-1865.

Holme, R. (2003). Drinking water contamination in Walkerton, Ontario: positive resolutions from a tragic event. *Water Sci Technol* **47,** 1-6.

Horcajo, P., Domínguez-Bernal, G., de la Fuente, R., Ruiz-Santa-Quiteria, J.A., Blanco, J.E., Blanco, M., *et al.* (2012). Comparison of ruminant and human attaching and effacing Escherichia coli (AEEC) strains. *Vet Microbiol* **155,** 341-348.

Huppertz, H.I., Rutkowski, S., Aleksic, S. and Karch, H. (1997). Acute and chronic diarrhoea and abdominal colic associated with enteroaggregative Escherichia coli in young children living in western Europe. *Lancet* **349,** 1660-1662.

Imamovic, L. and Muniesa, M. (2012). Characterizing RecA-independent induction of Shiga toxin2-encoding phages by EDTA treatment. *PLoS One* **7,** e32393.

Imamovic, L., Tozzoli, R., Michelacci, V., Minelli, F., Marziano, M.L., Caprioli, A. and Morabito, S. (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,** 4697-4704.

Iyoda, S., Tamura, K., Itoh, K., Izumiya, H., Ueno, N., Nagata, K.*, et al.* (2000). Inducible stx2 phages are lysogenized in the enteroaggregative and other phenotypic Escherichia coli O86:HNM isolated from patients. *FEMS Microbiol Lett* **191,** 7-10.

Jafari, A., Aslani, M.M. and Bouzari, S. (2012). Escherichia coli: a brief review of diarrheagenic pathotypes and their role in diarrheal diseases in Iran. *Iran J Microbiol* **4,** 102-117.

Jiang, Z.D., Greenberg, D., Nataro, J.P., Steffen, R. and DuPont, H.L. (2002). Rate of occurrence and pathogenic effect of enteroaggregative Escherichia coli virulence factors in international travelers. *J Clin Microbiol* **40,** 4185-4190.

Johnson, K.E., Thorpe, C.M. and Sears, C.L. (2006). The emerging clinical importance of non-O157 Shiga toxin-producing Escherichia coli. *Clin Infect Dis* **43,** 1587-1595.

Kaper (1996) Defining EPEC, pp. 130-133.

Kaper, J.B., Nataro, J.P. and Mobley, H.L. (2004). Pathogenic Escherichia coli. *Nat Rev Microbiol* **2,** 123-140.

Karch, H., Bielaszewska, M., Bitzan, M. and Schmidt, H. (1999). Epidemiology and diagnosis of Shiga toxin-producing Escherichia coli infections. *Diagn Microbiol Infect Dis* **34,** 229-243.

Karch, H., Denamur, E., Dobrindt, U., Finlay, B.B., Hengge, R., Johannes, L.*, et al.* (2012). The enemy within us: lessons from the 2011 European Escherichia coli O104:H4 outbreak. *EMBO Mol Med* **4,** 841-848.

Karch, H., Tarr, P.I. and Bielaszewska, M. (2005). Enterohaemorrhagic Escherichia coli in human medicine. *Int J Med Microbiol* **295,** 405-418.

Karmali, M.A. (1989). Infection by verocytotoxin-producing Escherichia coli. *Clin Microbiol Rev* **2,** 15-38.

Karmali, M.A., Mascarenhas, M., Shen, S., Ziebell, K., Johnson, S., Reid-Smith, R.*, 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,** 4930-4940.

Kim, J., Nietfeldt, J. and Benson, A.K. (1999). Octamer-based genome scanning distinguishes a unique subpopulation of Escherichia coli O157:H7 strains in cattle. *Proc Natl Acad Sci U S A* **96,** 13288-13293.

Kimmitt, P.T., Harwood, C.R. and Barer, M.R. (2000). Toxin gene expression by shiga toxin-producing Escherichia coli: the role of antibiotics and the bacterial SOS response. *Emerg Infect Dis* **6,** 458-465.

Klapproth, J.M., Scaletsky, I.C., McNamara, B.P., Lai, L.C., Malstrom, C., James, S.P. and Donnenberg, M.S. (2000). A large toxin from pathogenic Escherichia coli strains that inhibits lymphocyte activation. *Infect Immun* **68,** 2148-2155.

Käppeli, U., Hächler, H., Giezendanner, N., Beutin, L. and Stephan, R. (2011). Human infections with non-O157 Shiga toxin-producing Escherichia coli, Switzerland, 2000-2009. *Emerg Infect Dis* **17,** 180-185.

Laing, C.R., Zhang, Y., Gilmour, M.W., Allen, V., Johnson, R., Thomas, J.E. and Gannon, V.P. (2012). A comparison of Shiga-toxin 2 bacteriophage from classical enterohemorrhagic Escherichia coli serotypes and the German E. coli O104:H4 outbreak strain. *PLoS One* **7,** e37362.

Lan, R., Alles, M.C., Donohoe, K., Martinez, M.B. and Reeves, P.R. (2004). Molecular evolutionary relationships of enteroinvasive Escherichia coli and Shigella spp. *Infect Immun* **72,** 5080-5088.

Leopold, S.R., Dobrindt, U., Karch, H. and Mellmann, A. (2014) Genomic Plasticity and the Emergence of New Pathogenic Escherichia coli, in Pathogenic Escherichia coli, Molecular and Cellular Microbiology, Caister Academic Press.

Lima, I.F., Boisen, N., Quetz, J.a.S., Havt, A., de Carvalho, E.B., Soares, A.M.*, et al.* (2013). Prevalence of enteroaggregative Escherichia coli and its virulence-related genes in a case-control study among children from north-eastern Brazil. *J Med Microbiol* **62,** 683-693.

McDaniel, T.K., Jarvis, K.G., Donnenberg, M.S. and Kaper, J.B. (1995). A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc Natl Acad Sci U S A* **92,** 1664-1668.

Melton-Celsa, A.R. and O'Brien, A.D.i. (1998) *Escherichia coli* O157:H7 and Other Shiga Toxin-Producing E. coli Strains. ASM Press, Washington DC, USA, pp. 121-128.

Michelacci, V., Grande, L., Tozzoli, R., Maugliani, A., Caprioli, A. and Morabito, S. (2014). Identification of two allelic variants of toxB gene and investigation of their distribution among Verocytotoxin-producing Escherichia coli. *Int J Med Microbiol* **304,** 730-734.

Michelacci, V., Tozzoli, R., Caprioli, A., Martínez, R., Scheutz, F., Grande, L.*, et al.* (2013). 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* **19,** E149-156.

Morabito, S., Karch, H., Mariani-Kurkdjian, P., Schmidt, H., Minelli, F., Bingen, E. and Caprioli, A. (1998). Enteroaggregative, Shiga toxin-producing Escherichia coli O111:H2 associated with an outbreak of hemolytic-uremic syndrome. *J Clin Microbiol* **36,** 840-842.

Morabito, S., Tozzoli, R., Oswald, E. and 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,** 3343-3348.

Muniesa, M. and Schmidt, H. (2014) Shiga-toxin encoding Phages: Multifunctional Gene Ferries, in Pathogenic Escherichia coli, Molecular and Cellular Micriobiology, Caister Academic Press.

Nataro, J.P., Bopp, C.A., Fields, P.I., Kaper, J.B. and Strockbine, N.A. (2001) Escherichia, Shigella, and Salmonella. In Manual of Clinical Microbiology Washington D.C., USA: ASM Press.

Nataro, J.P. and Kaper, J.B. (1998a). Diarrheagenic Escherichia coli. *Clin Microbiol Rev* **11,** 142-201.

Nataro, J.P., Steiner, T. and Guerrant, R.L. (1998b). Enteroaggregative Escherichia coli. *Emerg Infect Dis* **4,** 251-261.

Nicholls, L., Grant, T.H. and Robins-Browne, R.M. (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.

Northey, G., Evans, M.R., Sarvotham, T.S., Thomas, D.R. and Howard, T.J. (2007). Sentinel surveillance for travellers' diarrhoea in primary care. *BMC Infect Dis* **7,** 126.

Ochman, H., Lawrence, J.G. and Groisman, E.A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* **405,** 299-304.

Ogura, Y., Ooka, T., Asadulghani, Terajima, J., Nougayrède, J.P., Kurokawa, K.*, et al.* (2007). Extensive genomic diversity and selective conservation of virulence-determinants in enterohemorrhagic Escherichia coli strains of O157 and non-O157 serotypes. *Genome Biol* **8,** R138.

Paton, A.W. and Paton, J.C. (1996). Enterobacter cloacae producing a Shiga-like toxin II-related cytotoxin associated with a case of hemolytic-uremic syndrome. *J Clin Microbiol* **34,** 463-465.

Paton, A.W., Srimanote, P., Woodrow, M.C. and Paton, J.C. (2001). Characterisation 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.

Paton, A.W., Woodrow, M.C., Doyle, R.M., Lanser, J.A. and Paton, J.C. (1999). Molecular characterisation of a Shiga toxigenic Escherichia coli O113:H21 strain lacking eae responsible for a cluster of cases of hemolytic-uremic syndrome. *J Clin Microbiol* **37,** 3357-3361.

Perna, N.T., Plunkett, G., Burland, V., Mau, B., Glasner, J.D., Rose, D.J.*, et al.* (2001). Genome sequence of enterohaemorrhagic Escherichia coli O157:H7. *Nature* **409,** 529-533.

Piérard, D., De Greve, H., Haesebrouck, F. and Mainil, J. (2012). O157:H7 and O104:H4 Vero/Shiga toxin-producing Escherichia coli outbreaks: respective role of cattle and humans. *Vet Res* **43,** 13.

Plunkett, G., Rose, D.J., Durfee, T.J. and Blattner, F.R. (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.

Qadri, F., Saha, A., Ahmed, T., Al Tarique, A., Begum, Y.A. and Svennerholm, A.M. (2007). Disease burden due to enterotoxigenic Escherichia coli in the first 2 years of life in an urban community in Bangladesh. *Infect Immun* **75,** 3961-3968.

Riley, L.W., Remis, R.S., Helgerson, S.D., McGee, H.B., Wells, J.G., Davis, B.R.*, et al.* (1983). Hemorrhagic colitis associated with a rare Escherichia coli serotype. *N Engl J Med* **308,** 681-685.

Russo, T.A. and Johnson, J.R. (2003). Medical and economic impact of extraintestinal infections due to Escherichia coli: focus on an increasingly important endemic problem. *Microbes Infect* **5,** 449-456.

Sansonetti, P. (2002). Host-pathogen interactions: the seduction of molecular cross talk. *Gut* **50 Suppl 3,** III2-8.

Scaletsky, I.C., Fabbricotti, S.H., Carvalho, R.L., Nunes, C.R., Maranhão, H.S., Morais, M.B. and Fagundes-Neto, U. (2002). Diffusely adherent Escherichia coli as a cause of acute diarrhea in young children in Northeast Brazil: a case-control study. *J Clin Microbiol* **40,** 645-648.

Sca*via*, G., Morabito, S., Tozzoli, R., Michelacci, V., Marziano, M.L., Minelli, F.*, et al.* (2011). Similarity of Shiga toxin-producing Escherichia coli O104:H4 strains from Italy and Germany. *Emerg Infect Dis* **17,** 1957-1958.

Scheutz, F., Nielsen, E.M., Frimodt-Møller, J., Boisen, N., Morabito, S., Tozzoli, R.*, 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**.

Scheutz, F., Teel, L.D., Beutin, L., Piérard, D., Buvens, G., Karch, H.*, et al.* (2012). Multicenter evaluation of a sequence-based protocol for subtyping Shiga toxins and standardizing Stx nomenclature. *J Clin Microbiol* **50,** 2951-2963.

Schmidt, H. (2001). Shiga-toxin-converting bacteriophages. *Res Microbiol* **152,** 687-695.

Schmidt, H., Beutin, L. and Karch, H. (1995). Molecular analysis of the plasmid-encoded hemolysin of Escherichia coli O157:H7 strain EDL 933. *Infect Immun* **63,** 1055-1061.

Schmidt, H., Bielaszewska, M. and Karch, H. (1999). Transduction of enteric Escherichia coli isolates with a derivative of Shiga toxin 2-encoding bacteriophage phi3538 isolated from Escherichia coli O157:H7. *Appl Environ Microbiol* **65,** 3855-3861.

Schmidt, H., Montag, M., Bockemühl, J., Heesemann, J. and Karch, H. (1993). Shiga-like toxin II-related cytotoxins in Citrobacter freundii strains from humans and beef samples. *Infect Immun* **61,** 534-543.

Smith, D.L., Rooks, D.J., Fogg, P.C., Darby, A.C., Thomson, N.R., McCarthy, A.J. and Allison, H.E. (2012). Comparative genomics of Shiga toxin encoding bacteriophages. *BMC Genomics* **13,** 311.

Smith, J.L., Fratamico, P.M. and Gunther, N.W. (2007). Extraintestinal pathogenic Escherichia coli. *Foodborne Pathog Dis* **4,** 134-163.

Steiner, T.S. (2014). The Worst of Both Worlds: Examining the Hypervirulence of the Shigatoxigenic/Enteroaggregative Escherichia coli O104:H4. *J Infect Dis*.

Stevens, M.P., Roe, A.J., Vlisidou, I., van Diemen, P.M., La Ragione, R.M., Best, A.*, et al.* (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.

Steyert, S.R., Sahl, J.W., Fraser, C.M., Teel, L.D., Scheutz, F. and Rasko, D.A. (2012). Comparative genomics and stx phage characterisation of LEE-negative Shiga toxin-producing Escherichia coli. *Front Cell Infect Microbiol* **2,** 133.

Strauch, E., Lurz, R. and Beutin, L. (2001). Characterisation of a Shiga toxin-encoding temperate bacteriophage of Shigella sonnei. *Infect Immun* **69,** 7588-7595.

Susskind, M.M. and Botstein, D. (1978) Molecular genetics of bacteriophage P22. In *Microbiol Rev.* pp. 385-413.

Sánchez, S., Beristain, X., Martínez, R., García, A., Martín, C., Vidal, D.*, et al.* (2012). Subtilase cytotoxin encoding genes are present in human, sheep and deer intimin-negative, Shiga toxin-producing Escherichia coli O128:H2. *Vet Microbiol* **159,** 531-535.

Tarr, C.L., Nelson, A.M., Beutin, L., Olsen, K.E. and Whittam, T.S. (2008). Molecular characterisation reveals similar virulence gene content in unrelated clonal groups of Escherichia coli of serogroup O174 (OX3). *J Bacteriol* **190,** 1344-1349.

Tatsuno, I., Kimura, H., Okutani, A., Kanamaru, K., Abe, H., Nagai, S.*, et al.* (2000). Isolation and characterisation of mini-Tn5Km2 insertion mutants of enterohemorrhagic Escherichia coli O157:H7 deficient in adherence to Caco-2 cells. *Infect Immun* **68,** 5943-5952.

Tenaillon, O., Skurnik, D., Picard, B. and Denamur, E. (2010). The population genetics of commensal Escherichia coli. *Nat Rev Microbiol* **8,** 207-217.

Tompkins, D.S., Hudson, M.J., Smith, H.R., Eglin, R.P., Wheeler, J.G., Brett, M.M.*, et al.* (1999). A study of infectious intestinal disease in England: microbiological findings in cases and controls. *Commun Dis Public Health* **2,** 108-113.

Tozzoli, R., Caprioli, A., Cappannella, S., Michelacci, V., Marziano, M.L. and Morabito, S. (2010). Production of the subtilase AB5 cytotoxin by Shiga toxin-negative Escherichia coli. *J Clin Microbiol* **48,** 178-183.

Tozzoli, R., Caprioli, A. and Morabito, S. (2005). Detection of toxB, a plasmid virulence gene of Escherichia coli O157, in enterohemorrhagic and enteropathogenic E. coli. *J Clin Microbiol* **43,** 4052-4056.

Tozzoli, R., Grande, L., Michelacci, V., Ranieri, P., Maugliani, A., Caprioli, A. and Stefano, M. (2014a). Shiga toxin-converting phages and the emergence of new pathogenic Escherichia coli: a world in motion. *Frontiers in Cellular and Infection Microbiology*.

Tozzoli, R. and Scheutz, F. (2014b) Diarrhoeagenic *Escherichia coli* infections in Humans, in Pathogenic Escherichia coli, Molecular and Cellular Microbiology, Caister Academic Press.

Tran Van Nhieu, G., Bourdet-Sicard, R., Duménil, G., Blocker, A. and Sansonetti, P.J. (2000). Bacterial signals and cell responses during Shigella entry into epithelial cells. *Cell Microbiol* **2,** 187-193.

Tschape, H., Prager, R., Streckel, W., Fruth, A., Tietze, E. and Böhme, G. (1995). Verotoxinogenic Citrobacter freundii associated with severe gastroenteritis and cases of haemolytic uraemic syndrome in a nursery school: green butter as the infection source. *Epidemiol Infect* **114,** 441-450.

Valentiner-Branth, P., Steinsland, H., Fischer, T.K., Perch, M., Scheutz, F., Dias, F.*, et al.* (2003). Cohort study of Guinean children: incidence, pathogenicity, conferred protection, and attributable risk for enteropathogens during the first 2 years of life. *J Clin Microbiol* **41,** 4238-4245.

van Elsas, J.D., Semenov, A.V., Costa, R. and Trevors, J.T. (2011). Survival of Escherichia coli in the environment: fundamental and public health aspects. *ISME J* **5,** 173-183.

van Hoek, A.H., Aarts, H.J., Bouw, E., van Overbeek, W.M. and Franz, E. (2013). The role of rpoS in Escherichia coli O157 manure-amended soil survival and distribution of allelic variations among bovine, food and clinical isolates. *FEMS Microbiol Lett* **338,** 18-23.

Venturini, C., Beatson, S.A., Djordjevic, S.P. and Walker, M.J. (2010). Multiple antibiotic resistance gene recruitment onto the enterohemorrhagic Escherichia coli virulence plasmid. *FASEB J* **24,** 1160-1166.

Vogeleer, P., Tremblay, Y.D., Mafu, A.A., Jacques, M. and Harel, J. (2014). Life on the outside: role of biofilms in environmental persistence of Shiga-toxin producing Escherichia coli. *Front Microbiol* **5,** 317.

Watanabe, Y., Ozasa, K., Mermin, J.H., Griffin, P.M., Masuda, K., Imashuku, S. and Sawada, T. (1999). Factory outbreak of Escherichia coli O157:H7 infection in Japan. *Emerg Infect Dis* **5,** 424-428.

Werts, C., Michel, V., Hofnung, M. and Charbit, A. (1994). Adsorption of bacteriophage lambda on the LamB protein of Escherichia coli K-12: point mutations in gene J of lambda responsible for extended host range. *J Bacteriol* **176,** 941-947.

Wester, A.L., Brandal, L.T. and Dahle, U.R. (2013). Extraintestinal pathogenic Escherichia coli carrying the Shiga Toxin gene stx2. *J Clin Microbiol* **51,** 4279-4280.

Wick, L.M., Qi, W., Lacher, D.W. and Whittam, T.S. (2005). Evolution of genomic content in the stepwise emergence of Escherichia coli O157:H7. *J Bacteriol* **187,** 1783-1791.

APPENDICES

APPENDIX 1:LIST OF ABBREVIATIONS

A/E: attaching-and-effacing
CF or CFA: colonization factor
CFU: colony-forming unit
CIEB: calf intestinal epithelial cell line
DAEC: diffusely adherent *E. coli*
DEC: diarrheagenic *E. coli*
bp: base pair
Da: dalton
DNA: deoxyribonucleic acid
EAggEC: enteroaggregative *E. coli*
EAHEC: enteroaggregative haemorrhagic *E. coli*
ECDC: european centre for disease prevention and control
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
MLST: multilocus sequence typing
NMEC: neonatal meningitis *E. coli*
ORF: open reading frame
PAI: pathogenicity island
PT: phage type
SPATE: serine protease autotransporters of Enterobacteriaceae
SPT: seropathotype
ST: heat-stable toxin
VT1: Verocytotoxin 1
VT2: Verocytotoxin 2
SEPEC: *sepsis*-associated *E. coli*
STEC: shiga toxin-producing *E. coli*
Tir: translocated intimin receptor
T3SS: type III secretion system
UPEC: uropathogenic *E. coli*
VTEC: verocytotoxin-producing *E. coli*
WGS: whole-genome sequencing

APPENDIX 2: LIST OF FIGURES