DOCTORAL SCHOOL IN BIOLOGY

Section: Biomedical Sciences and Technologies

CYCLE XXX

Shiga-toxin-producing *Escherichia coli* (STEC): Evolution of the different pathogenic clones and study of their interactions with the microbiota in the intestine and the environment

Federica Gigliucci

MSc. (Hons), Biology

Tutor

Dr. Stefano Morabito

Coordinator

Prof. Paolo Visca

Sommario

Gli *Escherichia coli* produttori di Shiga tossina (STEC) rappresentano un sottogruppo degli *E. coli* Diarreagenici (DEC), una popolazione eterogenea di *Escherichia coli* patogeni che comprende 6 differenti gruppi di *E. coli* coinvolti in patologie intestinali associate a diarrea e che, in aggiunta agli STEC, include *E. coli* Enteropatogeni (EPEC); *E. coli* Enterotossigenici (ETEC); *E. coli* Enteroinvasivi (EIEC); *E. coli* ad Aderenza Diffusa (DEAC); *E. coli* Enteroaggregativi (EAEC). Questi gruppi, o patotipi, differiscono tra loro per il meccanismo di virulenza e per le manifestazioni cliniche indotte.

Gli STEC sono microorganismi zoonotici e i ruminanti rappresentano il loro naturale serbatoio. Essi costituiscono un eterogeneo gruppo di *E. coli* patogeni, caratterizzati dalla capacità di produrre una potente citotossina, definita Shiga tossina (Stx), acquisita in seguito all'infezione di batteriofagi portatori dei geni *stx*. Le Stx vengono classificate in due principali varianti antigenicamente diverse, Stx1 (gene *stx1*) e Stx2 (gene *stx2*), delle quali sono noti diversi sottotipi, rispettivamente, Stx1a, Stx1c, Stx1d e Stx2a-Stx2g. Gli STEC, oltre ai geni *stx* codificanti le Shiga tossine, possiedono ulteriori fattori di virulenza, che conferiscono loro la capacità di colonizzare efficientemente la mucosa intestinale nell'ospite umano, causando malattie anche gravi, tra cui la colite emorragica e la sindrome emolitico uremica (SEU). Tali geni di virulenza sono tutti veicolati da Elementi Genetici Mobili (MGE), i quali possono essere facilmente acquisiti o persi dai diversi ceppi di *E. coli* patogeni, grazie alla elevata plasticità del loro genoma.

Il principale obiettivo di questo lavoro è stato quello di investigare i determinanti di esposizione dell'uomo alle infezioni da STEC, attraverso l'identificazione di nuovi serbatoi animali e veicoli di diffusione delle infezioni al fine di comprenderne i cicli epidemiologici. E' stato inoltre obiettivo di questo lavoro lo studio dell'interazione di ceppi STEC con la microflora presente in veicoli di matrice ambientale e nell'intestino dell'ospite umano.

Benché i ruminanti siano considerati il serbatoio naturale degli STEC, questi microrganismi sono stati isolati da numerose altre specie animali, che agiscono come ospiti transitori o permanenti. I piccioni, in particolare, sono considerati il *reservoir* di un sottogruppo di STEC, gli *E. coli* produttori di Stx2f, i quali sono stati a lungo considerati di scarsa rilevanza in sanità pubblica, in quanto isolati raramente dall'uomo e da casi di diarrea non complicata. Tuttavia, negli ultimi anni, sporadici casi di SEU sono stati associati a STEC produttori di Stx2f. Tramite sequenziamento dell'intero genoma (WGS, whole-genome sequencing) abbiamo esaminato e comparato ceppi di STEC-Stx2f isolati da piccioni sani, da casi di diarrea non complicata e da casi di SEU. L'analisi ha indicato l'esistenza di tre distinte sottopopolazioni di *E. coli* produttori di Stx2f sulla base di una differente distribuzione di fattori di virulenza. I ceppi isolati dalle feci di pazienti con SEU

mostravano la presenza dell'isola di patogenicità OI-122 e del plasmide di virulenza pO157, normalmente associati alle forme più gravi dell'infezione da STEC, che non erano mai stati descritti prima in STEC portatori del gene *stx2f*. I nostri risultati hanno dimostrato che la Shiga tossina codificata dai geni *stx2f* presenti nei batteriofagi integrati nei ceppi STEC del piccione è in grado di causare la SEU se prodotta nel contesto genomico adeguato e suggeriscono che questa specie animale possa essere considerata serbatoio di ceppi STEC altamente virulenti o di batteriofagi-*stx2f* in grado di disperdersi e lisogenizzare ceppi di *E. coli* con un corredo di geni di colonizzazione adeguato, quali quelli presenti in diverse specie animali e nell'ambiente, rappresentando un serio problema di sanità pubblica.

L'ambiente, in particolare quello destinato all'allevamento, è considerato un'importante sorgente delle infezioni per la presenza degli animali serbatoio che eliminano i ceppi STEC.

Per migliorare la qualità del terreno, intensificare la produzione agricola e ridurre i costi legati alla fertilizzazione dei suoli con ammendanti chimici, si sta intensificando l'utilizzo dei cosiddetti "Biosolidi" (BSO). Con questo termine vengono definiti ammendanti composti da rifiuti di natura organica provenienti da aziende agricole e zootecniche, fanghi di depurazione civile, e rifiuti di origine urbana, i quali vengono sottoposti a specifici trattamenti prima di venir applicati sul suolo agricolo. A fronte della scarsità dei dati presenti in letteratura circa il potenziale di queste matrici di veicolare agenti di zoonosi, abbiamo sottoposto a screening molecolare un set di campioni di biosolidi. Attraverso questo approccio abbiamo osservato nei campioni la presenza di tratti di virulenza associati a Norovirus, Adenovirus e a i principali patogruppi di *E. coli* diarreagenici. In particolare, una buona parte dei BSO analizzati è risultata positiva per la presenza dei geni *stx*. Inoltre, i risultati dei nostri esperimenti suggeriscono la presenza di microorganismi vitali nei campioni analizzati.

Dal momento che questo studio preliminare aveva evidenziato come i BSO possano rappresentare un veicolo di *E. coli* patogeni, inclusi gli STEC, abbiamo effettuato una caratterizzazione completa, non selettiva, dei possibili pericoli microbiologici associati al loro utilizzo. Abbiamo sviluppato un approccio di metagenomica che ci ha permesso sia di identificare importanti determinanti di virulenza, nonché di ottenere un profilo tassonomico delle comunità microbiche presenti nei differenti BSO analizzati. L'analisi ha mostrato la presenza di geni di virulenza associati a vari patogruppi di *E. coli* diarreagenici, confermando i risultati precedentemente ottenuti nello screening molecolare per PCR. Inoltre, ha identificato la presenza di tratti genetici conferenti capacità di resistenza ai metalli pesanti e agli antibiotici.

L’analisi metagenomica si è rilevata uno strumento efficace nella caratterizzazione diretta di comunità microbiche all’interno di matrici complesse, superando i classici metodi di coltura e isolamento batterico.

Lo stesso approccio è stato anche utilizzato per esaminare possibili cambiamenti nella composizione del microbiota intestinale in campioni fecali di pazienti con infezione da STEC, in comparazione a quello di individui sani. Lo studio è stato effettuato analizzando campioni prelevati nel corso di un’epidemia di infezione da STEC di sierogrupo O26 (STEC O26) e mettendo a punto due differenti approcci di analisi bioinformatica, che hanno consentito di ottenere una profonda classificazione tassonomica dei vari campioni fecali analizzati. L’analisi ha confermato l’ipotesi di una perturbazione non casuale della microflora intestinale conseguente all’infezione da STEC. In dettaglio, i campioni fecali di pazienti malati presentavano una minore abbondanza di specie commensali appartenenti agli ordini *Bifidobacteriales* e *Costridiales* rispetto ai campioni prelevati di individui sani o guariti dall’infezione. Il nostro studio rappresenta il primo tentativo di ottenere evidenze sulle alterazioni del microbiota intestinale nel corso di infezioni da STEC e ha consentito di identificare alcune delle dinamiche che si instaurano tra le popolazioni microbiche dell’ambiente intestinale, confermando quanto descritto in altri studi circa il ruolo benefico di varie specie appartenenti agli ordini *Bifidobacteriales* e *Clostridiales*, sia nell’interferire con la colonizzazione gastrointestinale di enteropatogeni, sia di stimolare la risposta immunitaria. Infine, abbiamo effettuato nei metagenomi la ricerca dei geni di virulenza degli *E. coli* patogeni. Questo studio ha permesso di identificare il viruloma completo del ceppo epidemico STEC O26, precedentemente determinato tramite sequenziamento dell’intero genoma.

I risultati ottenuti hanno mostrato come la metagenomica rappresenti uno strumento promettente per la diagnosi clinica delle malattie infettive, superando la necessità dei metodi classici basati sull’isolamento batterico.

Il lavoro descritto in questa tesi si propone come uno studio effettuato sull’intero asse di esposizione dell’uomo alle infezioni da STEC, dagli animali serbatoio all’influenza sulla flora intestinale umana dei ceppi STEC in seguito a infezione. I nostri risultati hanno consentito di identificare nuovi cicli di trasmissione delle infezioni attraverso la descrizione di animali serbatoio precedentemente non associati alle infezioni più gravi e di nuovi possibili veicoli ambientali per la disseminazione delle infezioni. Infine, l’utilizzo di approcci tecnologici avanzati ha consentito di derivare le prime informazioni sui cambiamenti dinamici delle popolazioni microbiche intestinali in seguito all’infezione.

Table of contents

Chapter 1

Introduction

1. Introduction

1.1 *Escherichia coli* (*E. coli*)

Escherichia coli is a facultative anaerobic, gram-negative bacterium belonging to the family of *Enterobacteriaceae* that inhabits the intestinal tract of human and warm-blooded animals. *E. coli* was first described by the German-Austrian paediatrician Theodor Escherich in 1885, as *Bacterium coli commune*, which he isolated from the faeces of healthy infants (Escherich, 1989). This bacterial species typically colonizes the gastrointestinal tract during the first phases of life, establishing mutual beneficial relationships with the host and playing an important role in maintaining the equilibrium between the numerous bacterial species colonising the gut.

E. coli is one of the bacterial species with the widest adaptability to a broad range of niches either within organisms, because some strains can colonize a range of tissues and organs other than the intestine, or environmental niches including water and soils. The capability to colonize so many hosts, body compartments and environmental niches is due to their high ability to acquire and loose DNA sequences from different bacterial species. This remarkable genomic plasticity characterizing the genome of *E. coli* is the basis of its huge diversity and adaptive gene content.

The exchange of DNA traits between bacteria is defined as horizontal gene transfer (HGT) and is characterized by the integration and excision of mobile genetic elements (MGE). The MGEs are segments of DNA, including plasmids, genomic islands, bacteriophages, integrons, transposons and insertion sequences, encoding all the components that mediate the movement of DNA within genomes (intracellular mobility) or between bacterial cells (intercellular mobility) (Frost *et al.*, 2005).

The HGT contributed to the emergence of pathogenic variants of *E. coli* from possible commensal ancestors and allowed the adaptation of the pathogenic types to evade host immune defences (Croxen *et al.*, 2010).

The pathogenic *E. coli* can cause a wide spectrum of diseases in humans and are divided in two main pathogroups, defined on the basis of which part of the organism they colonize: Extraintestinal Pathogenic *E. coli* (ExPEC) and Diarrheagenic *E. coli* (DEC).

ExPEC usually possesses diverse virulence features conferring the capability to colonize host districts outside the gastrointestinal tract, including blood, lungs, central nervous system, and urinary tract, resulting in disease (Russo *et al.*, 2000, Smith *et al.*, 2007). This pathogroup is mainly divided in two sub-populations: Uropathogenic *E. coli* (UPEC), primary cause of the vast majority of urinary tract infections worldwide (Wiles *et al.*, 2008, Bien *et al.*, 2012) and Neonatal Meningitis

*E. co*li (NMEC), considered one of the predominant agents of neonatal bacterial meningitis (Nicholson *et al.*, 2016).

Diarrheagenic *E. coli* (DEC) is the biggest group of pathogenic *E. coli* comprising strains involved in diarrheal diseases. The DEC pathogroup includes six different pathotypes, which differ for virulence mechanisms and consequently for clinical symptoms in the patients, according to the set of virulence determinants acquired. The DEC pathotypes are in fact classified as Enteropathogenic *E. coli* (EPEC), Shiga toxin-producing *E. coli* (STEC), Enteroaggregative *E. coli* (EAEC), Enterotoxigenic *E. coli* (ETEC), Enteroinvasive *E. coli* (EIEC) and Diffusely Adherent *E. coli* (DAEC) (Nataro *et al.*, 1998).

Enteropathogenic E. coli (EPEC)

EPEC is the first DEC pathotype described, after the first asocciation of a *E. coli* strain with outbreaks of infantile diarrhoea in 1940 (Bray, 1945, Kaper *et al.*, 2004). The pathogenesis of EPEC depends on their ability to efficiently adhere to human intestinal mucosa, by means of bundle-forming pili, and to cause the "attaching and effacing" (A/E) lesion to the enterocytes. Such mechanism is governed by virulence factors present in the pathogenicity island (PAI) defined as the Locus of Enterocyte Effacement (LEE) (Donnenberg *et al.*, 2001, Kaper *et al.*, 2004) (described below in detail). The LEE-PAI was later on identified also in some STEC strains (Karmali *et al.*, 2003).

The EPEC group is dividided in two subgroup, the typical EPEC (tEPEC) and atypical EPEC (aEPEC). The tEPEC harbour a large virulence plasmid called EPEC Adherence Factor (EAF), which carries different genes encoding the bundle forming pili (*bfp* locus) as well as genes involved in the regulation of A/E mechanism (Baldini *et al.*, 1983). The tEPEC infections occur through inter-human contacts and oral-faecal routes. On the contrary, the aEPEC lack the EAF plasmid, but possess a large virulence plasmid carrying additional virulence factors described also in STEC pathogroup (see below), including the enterohaemolysin (*ehx* gene). Similarly to STEC, aEPEC are zoonotic pathogens and may cause foodborne oubreaks.

Enteroaggregative E. coli (EAEC)

The *E. coli* strains belonging to EAEC pathogroup adhere to the terminal ileum and colon producing a peculiar "stacked-brick" pattern of aggregation, defined aggregative adherence (AA) (Nataro *et al.*, 1987). This mechanism is mediated by aggregative adhrence fimbrie (AAF/I to AAF/V), whose genes are conveyed on a 55-65 MDa plasmids (pAAs) (Vial *et al.*, 1988). Such plasmid also carries the *aap* gene encoding dispersin, an antigenic antiaggregative protein, which

favours the adhesion process; the *aat* operon encoding the ABC transporter responsible for dispersin secretion; and other virulence determinants encoding toxins which play an important role in causing the diarrhoea, such as the serine protease autotransporters of Enterobateriaceae (SPATEs), the EAEC heat-stable cytotoxin 1 (EAST-1) and Shigella enterotoxin-1 (ShET1) (Hebbelstrup Jensen *et al.*, 2014). The expression of all these virulence factors is regulated by the transcriptional regulator Aggr.

EAEC infections occurs through person-to-person transmission and through ingestion of contaminated food and water. The typical clinical symptoms associated with EAEC infection are watery and mucoid persistent diarrhoea, with or without blood, vomiting and low-grade fever (Adachi *et al.*, 2002).

Enterotoxigenic E. coli (ETEC)

ETEC are responsible in humans for an acute form of secretory watery diarrhoea in children and travelers from industrialized to developing countries (Northey *et al.*, 2007). ETEC strains can colonize the gastrointestinal tract, through the expression of adhesins, wich attach specific receptors on enterocytes of the intestinal lumen. To date, more than twenty variants of these adhesins have been described, defined as colonization factors (CFs or CFAs) based on different types of structures (fimbriae, fibres and non-fimbrial) (Del Canto *et al.*, 2014). The pathogenicity of ETEC is characterized also by the production of heat-stable (ST) or heat-labile (LT) enterotoxins, encoded by *st* and *lt* genes respectively and carried by plasmids (Smith *et al.*, 1970).

ETEC can colonize both animal and human intestines. Indeed, different species-specific variants of enterotoxins have been found in ETEC strains, as the STh described in human host and STp associated with disease in pigs, as well as different CFAs conferring host specificity. Human infection by ETEC mainly occur through ingestion of contaminated food, especially vegetables, and water.

Enteroinvasive E.coli (EIEC)

EIEC strains were initially classified as a *Shigella* spp., because of their close genetic and phenotypic relation, and then considered a new pathogroup of pathogenic *E. coli*. In particular, EIEC and *Shigell*a strains share a large virulence plasmid (pInv), that confer the capability to invade and destroy epithelial cells of the distal large bowel (Lan *et al.*, 2004). Such plasmid harbours *ipaA* to *ipaH* virulence antigens-coding genes, which mediate the invasion mechanism, and *mxi* and *spa* genes, encoding a type III secration system. EIEC infection may cause invasive infiammatory colitis and the typical DEC watery diarhoea, and in some cases bloody diarrhoea (Nataro *et al.*,

1998). The faecal-oral route is the main transmission pathway of EIEC infection, by the direct contact with infected humans, although transmission of the infections through food vehicles has been also described (Escher *et al.*, 2014).

Diffusely Adherent E. coli (DAEC)

DAEC are defined by a pattern of diffuse adherence (DA) on epithelial cells, in which the bacteria can cover the whole cell surface. This mechanism is mediated by a group of fimbial and afimbrial (Afa) adhesins, collectively defined as Afa/Dr family (Mansan-Almeida *et al.*, 2013).

The implication of DAEC strains in diarrhhoeal disease is still unclear, because DAEC strain were found in children with and without diarrhea. Nevertheless, several studies associated DAEC with disease in children elder than 12 months and in adults (Scaletsky *et al.*, 2002, Opintan *et al.*, 2010).

1.2 Shiga toxin-producing *E. coli* (STEC)

Shiga toxin-producing *E. coli* represents a population of foodborne pathogens distributed worldwide. STEC are strains of *E. coli* able to produce potent cytotoxins, the Shiga toxins (Stxs).

As STEC represent the focus of the work described in the present thesis, this DEC pathotype will be treated in more detail than the other DEC pathotypes.

STEC may cause severe afflictions in human, including haemorrhagic colitis (HC) and haemolytic uremic syndrome (HUS) (Karmali, 1989). The first evidence of STEC infection in humans dates back to 1983 when the Centers for Disease Control, Atlanta, Ga. linked two oubreaks of HC with a previously rarely isolated *E. coli* serotype, O157:H7. The outbreaks occurred in USA were associated with the ingestion of undercoked beef meat at a fast-food restaurant chain (Karmali *et al.*, 1983, Riley *et al.*, 1983). In the same period, several studies based on the isolation of STEC strains from faecal cultures and the demonstration of free Stxs in faeces from patients led to the associations between STEC and the haemolytic uremic syndrome (Karmali *et al.*, 1983, Karmali *et al.*, 1985). This syndrome is defined by acute renal failure, thrombocytopenia, and microangiopathic hemolytic anemia, features that develop few days after an acute, bloody diarrheal illness typical of hemorrhagic colitis. Several studies conducted during initial investigation of outbreaks and sporadic cases of HUS showed that STEC isolates belonged to O157:H7 and other several serotypes of *E. coli* (Karmali *et al.*, 1983, Karmali *et al.*, 1985). However, the spectrum of STEC types is continuously growing and nowadays includes dozens of different serogroups and serotypes (EUSR, 2016).

1.2.1 Main virulence features of STEC

Shiga toxins are the most important virulence factors of STEC. These were first described in 1977 by Konowalchuk and colleagues, who observed a profound cytopatic effect of the toxins produced by a group of *E. coli* strains isolated from diarrheal illness, onto Vero cells (kidney epithelial cells extracted from an African green monkey) culture (Konowalchuk *et al.*, 1977). The toxin was then named Vero cytotoxin (VT). Further studies demonstrated that the VT was structurally and functionally identical to the Shiga toxin produced by strains of *Shigella dysenteriae* serotype 1 (O'Brien *et al.*, 1983, Strockbine *et al.*, 1988). Afterwards, the Verocytotoxin assumed also the name of Shiga-like toxin, eventually defined as Shiga toxin and the bacteria able to produce it were termed Shiga toxin-producing *E. coli*.

Shiga toxin is a holotoxin of about 70 kDa composed of a single 32 kDa A subunit (StxA), responsible for the toxic activity of the protein, and a pentamer of five identical 7.7 kDa B subunits (StxB) (O'Brien *et al.*, 1987). The B subunit binds to a specific glycolipid receptor, globotriaosylceramide (Gb3), which is present on the surface of target cells (17). After binding, the toxin is endocytosed and transported to the endoplasmic reticulum, where the A subunit is proteolytically cleaved and then activated. StxA exhibits an RNA N-glycosidase activity against the 28S rRNA, that inhibits host protein synthesis and induces apoptosis (Sandvig *et al.*, 1992, Sandvig, 2001).

Stxs are classified in two major antigenically different types, Stx1 and Stx2, encoded by *stx1* and *stx2* genes. The Stx-encoding genes are harboured by lamboid bacteriophages, stably integrated in the bacterial chromosome in a lysogenic state (O'Brien *et al.*, 1984). The activation of lytic cycle induces the phage to infect new hosts, increasing the production of the toxin itself (Plunkett *et al.*, 1999).

Stx1 and Stx2 are structurally similar, sharing about 60% amino acid identity (Jackson *et al.*, 1987), but they differ in cytotoxicity. Shiga toxin 2, indeed, is more cytotoxic than Stx1 and it is most commonly associated with HUS (Ethelberg *et al.*, 2004, Basu *et al.*, 2015). Besides the two main variants, three subtypes of Stx1 (Stx1a, Stx1c and Stx1d) and seven of Stx2 (Stx2a to Stx2g) have been described (Persson *et al.*, 2007). STEC can possess either Stx1 or Stx2 alone or both and in different subtype combinations (Friedrich *et al.*, 2002, Persson *et al.*, 2007).

Besides the capability to produce the Shiga toxin, some STEC are able to colonise the intestinal mucosa of the host through the “attaching and effacing” (A/E) apparatus, a colonization mechanism shared with another DEC pathotype, EPEC (see above). It is commonly accepted that the STEC strains causing the A/E lesion are associated with the most severe forms of human infections including HUS. This mechanism generates the typical A/E lesion, where microvilli on enterocytes

collapse and are replaced by microfilamentous structures, known as pedestals, that protrude from the host cell surface, allowig an intimate adherence of the bacteria to the enterocyte plasma membrane (Moon *et al.*, 1983). As seen in EPEC, the A/E mechanism is mediated by various proteins that are encoded in a chromosomal pathogenicity island (PAI) defined as the Locus of Enterocyte Effacement (LEE) (Donnenberg *et al.*, 2001). The LEE locus encodes an adhesin called intimin (*eae* gene) and its traslocated receptor (Tir), which mediate the firm attachment of the bacterium to the host cell surface (Jerse *et al.*, 1990, Kenny *et al.*, 1997). It also contains genes encoding for type III secretion system (T3SS), including translocator proteins such as EspA, EspB, and EspD (Hartland *et al.*, 2000) that export effector proteins to the host cell, as well as genes encoding the effectors EspF, EspG, EspZ, EspH, and Map (Wong *et al.*, 2011). Additionally, non-LEE encoded effectors (nle) proteins have been described in LEE-positive STEC, which are traslocated by the T3SS and are known to be involved in the pathogenesis of STEC infections. These included NleA/EspI, NleB, NleC, NleD, NleE, EspJ, NleH, EspG, EspM, and Cif (Garmendia *et al.*, 2005).

Additional to the Stxs and the LEE-encoded proteins, other virulence factors, carried by several mobile genetic elements, have been described in different STEC strains responsible of human severe disease (Table 1).

Table 1. Some additional mobile genetics elements described to confer virulence to STEC

MGE	Virulence genes hosted	Features of related encoded proteins	Reference
PAI OI-122	*efa1/lifA*	Inibition of host lymphocyte activation; involvment in A/E lesion	(Klapproth *et al.*, 2000)
pO157	*ehxA*	Enteroemolysin, responsible of the lysis of enterocytes	(Beutin *et al.*, 1990)
	katP	Catalase-peroxidase, with bifunctional catalase and peroxidase activity	(Brunder *et al.*, 1996)
	espP	Serine-protease, with proteolytic activity against coagulation factor V	(Brunder *et al.*, 1997)
	stcE	Metalloprotease, interferes with the activation of the inflammatory pathway	(Grys *et al.*, 2005)
	toxB	Homologue of efa1/lifA favours A/E mechanism	(Tatsuno *et al.*, 2001)
pO113	*epeA*	Autotransporter	(Leyton *et al.*, 2003)
	saa	Autoagglutinating adhesin	(Paton *et al.*, 2001)
	subAb	Subtilase cytotoxin	(Paton *et al.*, 2004)

1.3 STEC infections in humans

STEC constituite a heterogeneous pathogroup of *E. coli* capable to cause a wide spectrum of human diseases upon infection. The clinical signs range from mild diarrhoea, haemorrhagic colitis (HC) to the haemolitic uremic syndrome (HUS). HUS is a life-threatening syndrome, representing the most severe form of STEC infection. Usually, it occurs in children, elders and immuno-compromised patients (Ochoa *et al.*, 2003), but the appearance of new STEC strains with mixed virulence features, following the acquisition of Stx-converting bacteriophages by different *E. coli* pathotypes, may involve in severe diseases in healthy adults (Tozzoli *et al.*, 2014).

Besides the pathogenic potential of the infecting strains, the age and immune status of patients, also the human intestinal microbiota appears to be involved in the progression of the infections towards the more severe forms. It has been proposed that intestinal microflora can act as amplifier of the

Stx-converting phages resulting in augmented ability to produce the toxin (Gamage *et al.*, 2003). Intestinal *E. coli*, after the acquisition of *stx*-phages and their cell lysis, released both phages and toxins, resulting in higher level of Stxs in the intestine. On the contrary, the presence of commensal *E. coli* resistant to the phage infection, resulted in reduced amount of Stxs, because its production was limitated only to the original infecting STEC strain (Gamage *et al.*, 2003). Therefore, the authors suggested that the presence of intestinal microflora susceptible to *stx*-phages may influence the severity of disease through this mechanism of phage amplification.

On the other hand, several studies proposed that bacterial species composing the gut microbiota can interfere with the gastrointestinal pathogens colonization. It has been shown that commensal bacterial species competed with enterotoxigenic *E. coli* (ETEC) strains for common binding sites, inhibiting their adherence to human intestinal epithelial cell (Fujiwara *et al.*, 2001). Moreover, a preliminary study demonstrated the capability of some *Bifidobacterium* strains, livinig in the human gastrointestinal tract, to exclude or limit the presence of enteropathogens in this district (Gueimonde *et al.*, 2007). The proposed model is based on the augmented adhesion of *Bifidobacterium* spp. to human intestinal mucus, inhibiting the colonization of some pathogens and displacing those previously adhered to mucus (Gueimonde *et al.*, 2007). Additionally, many intestinal bacteria may inhibit enteropathogens by competing for nutrients or by producing inhibitory substances (Figure 1). As an example, *Bacteroidetes thetaiotaomicron* competitively excludes *Citrobacter rodentium* from the intestine of mice used in the experiments, because of the consumption of specific carbohydrates (Kamada *et al.*, 2012). Moreover, several species normally present in the human intestinal microbiota produce butyrate and acetate, which act against pathogenic *E. coli* infection (Shin *et al.*, 2002), or alternatively they secrete bacteriocin, that inhibit the growth of enteric pahogens (Rolhion *et al.*, 2016) (Figure 1).

Even if the gut microflora seems to have a central role in the resistance to pathogens colonization, some pathogenic microorganisms have evolved strategies to overcome the microbiota actions (Rolhion *et al.*, 2016) (Figure 2). Such mechanisms include: the use of alternative nutrients or intestinal niches; promotion of host infiammation, through which the growth of commensal bacteria appeared inhibited; toxin production.

On the other hand, commensal bacteria may favour the establishment of infections with pathogens and even trigger their virulence mechanisms (Figure 2). As a matter of fact, intestinal microflora provide some metabolities which may enhance pathogens' virulence expression and their colonization of the host gut. It has been described that *Clostridium difficile* colonization is augmented by the production of succinate from *Bacteroidetes thetaiotaomicron* (Ferreyra et al.,

2014). Succinate can also regulate the expression of the locus of enterocyte effacement (LEE) genes in pathogenic *E. coli*, STEC and EPEC, contributing to their virulence (Curtis *et al.*, 2014).

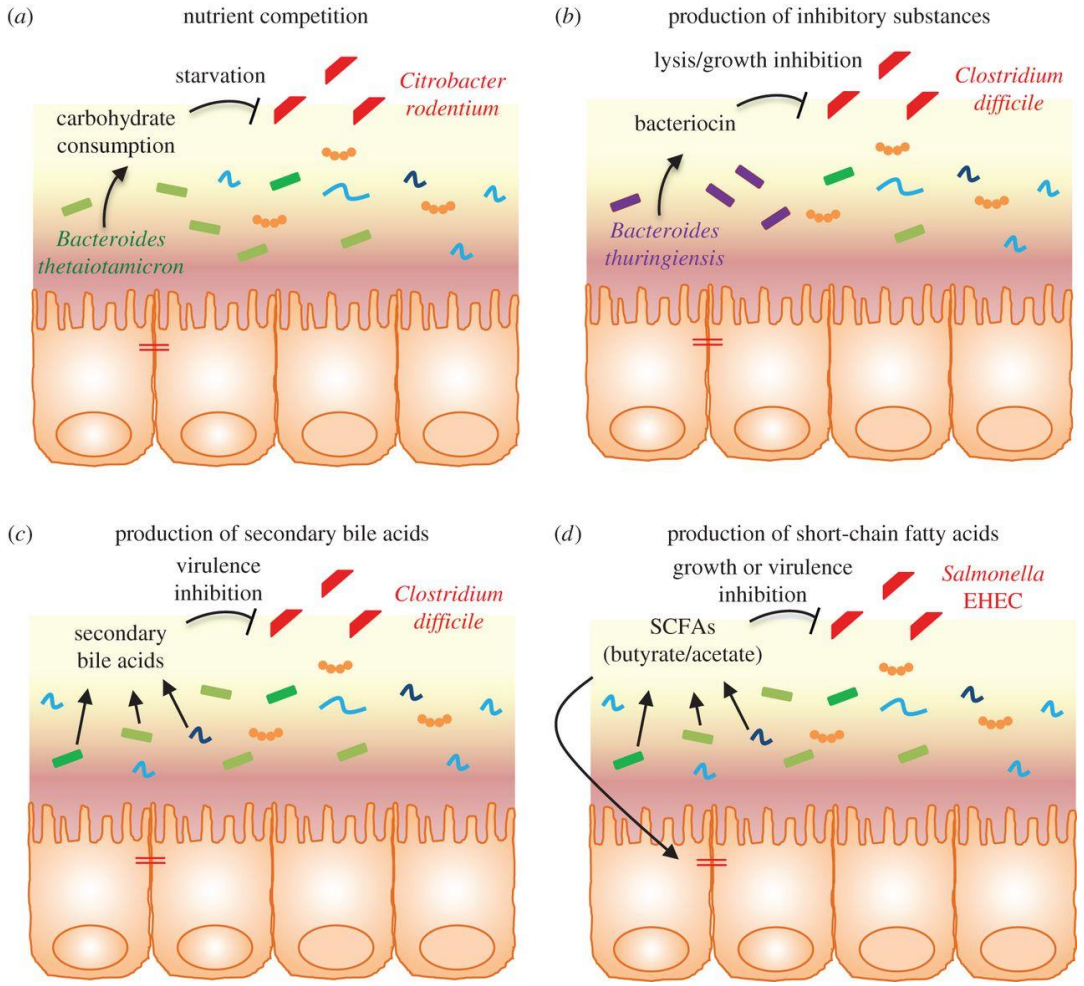


Figure 1.1 Role of the intestinal microbiota in colonization resistance by enteric pathogens (Rolhion *et al.*, 2016).

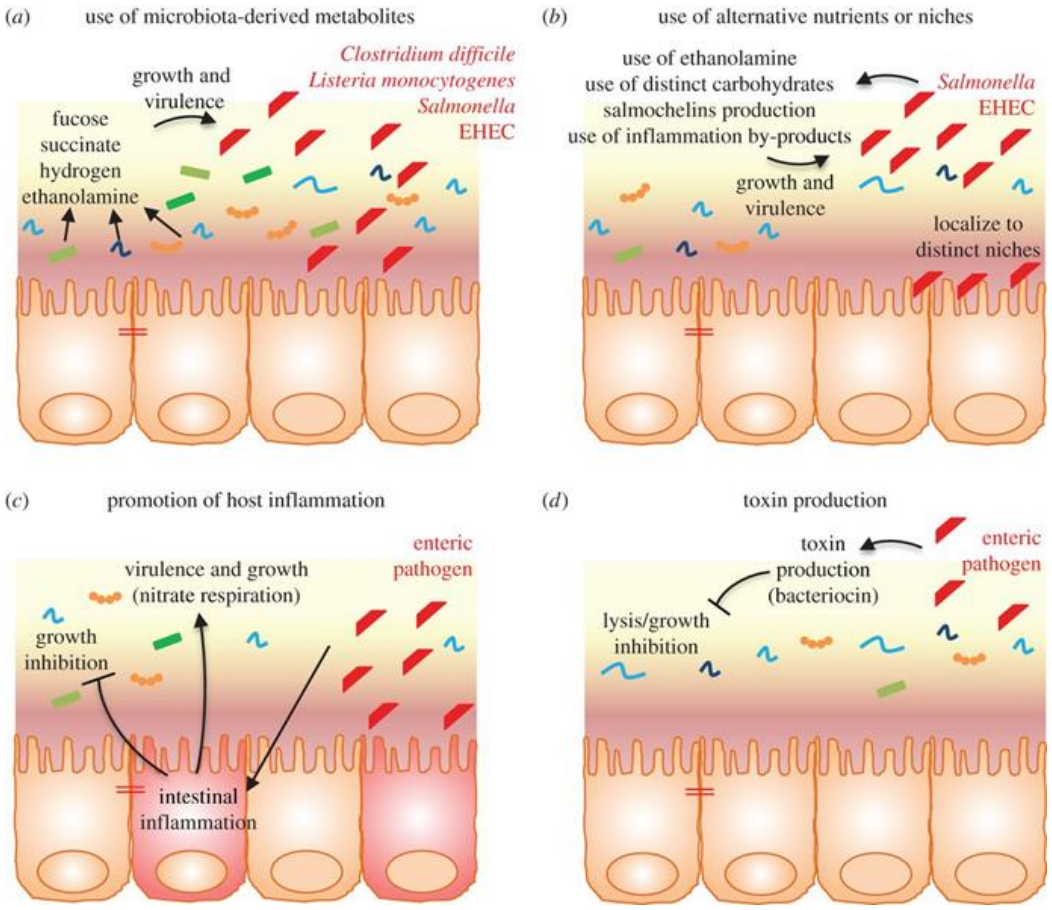


Figure 2. Mechanisms developed by enteric pathogens in order to exceed the intestinal community activities involved in the colonization resistance (Rolhion *et al.*, 2016).

It is well known that intestinal microorganisms have a profound impact on human physiology, being also able to modulate the local and systemic immune responses (Shreiner *et al.*, 2015). Perturbation of this equilibrium, through advanced age, use of antibiotics, or for the presence of infiammatory bowel diseases, alter the structure and function of intestinal micriobial community, with consequent dysbiosis and augmented susceptibility to colonization by enteric pathogens. It has been showed that changes in microbiota composition increased the probability of developing *Clostridium difficile* infection (CDI) (Lagier, 2016). Patients presenting CDI showed a dicreased biodiversity of microbiota, especially a reduced abundance of beneficial species (Lagier, 2016). Studies conducted on the alteration of microbiota composition upon *Clostridium difficile* infection led to the tratment of patients with the fecal microbiota transplant (FMT), in order to recovery the microbial diversity and to re-establish the intestinal predominance by beneficial bacteria (Kassam *et al.*, 2013).

1.4 Natural reservoir and other carriers of STEC

STEC are zoonotic phatogens able to colonize the gastrointestinal tract of various animal species. Ruminants, especially the cattle, are considered the main natural reservoir of STEC, because the intestinal colonization is asymptomatic for the absence of vascular receptors for Stxs (Pruimboom-Brees *et al.*, 2000). Other ruminants, including sheep, deer, goat, act as reservoir and contribute to the spread of STEC in the environment, excreting the bacteria through their feces. Consequently, other animals, as swine, dogs, horses, rabbits and birds, may be in contact with STEC and can act as spillover host. These animals may be susceptible or not to the STEC colonization, but can be transitory host the bacteria and disseminate them, increasing the risk to transmit the pathogens to humans (Caprioli *et al.*, 2005).

In the past, several avian species were identified as potential source of STEC infection (Wallace *et al.*, 1997, Nielsen *et al.*, 2004), with some of them, as the pigeons, living in close association with the human population (Morabito *et al.*, 2001). Over the years, pigeons have acquired more and more importance in public health, being a natural reservoir of STEC strains producing a non common variant of Shiga toxin type 2, Stx2f, (Morabito *et al.*, 2001).

The first association between pigeons and STEC carrying the *stx2f* gene was in 1998, when Dell'Omo and colleagues described the presence of STEC in the faeces of these birds (Dell'Omo *et al.*, 1998). Later on, the Stx-coding genes were characterized and the Stx2f variant of the toxin was described (Schmidt *et al.*, 2000). At that time, Stx2f-producing *E. coli* were considered only adapted to pigeons, with a limited impact on human illness. However, the number of Stx2f-STEC strains isolated from cases of human infection is increased in the last years. Several cases of mild diarrhoea were reported in Germany, from 2004 to 2007 (Prager *et al.*, 2009), and in the Netherland from 2008 to 2011 (Friesema *et al.*, 2014) as well as an HUS case was associated with the presence of a Stx2f-producing O8:H19 *E. coli* strain in the Netherland (Friesema *et al.*, 2015). Furthermore, STEC strains isolated from pigeons usually possess the LEE locus, representing in principle, a serious risk for human health.

1.5 Transmission pathways and environmental contamination

The main transmission route of Shiga toxin-producing *E. coli* to humans is the ingestion of contaminated food of animal origin, vegetables or drinking water, but also through inter-human cycles or direct contacts with STEC positive animals, or their environment. Indeed, faecal shedding of STEC from asymptomatic or symptomatic animals involves a wide contamination of farmland, water and consequently fresh produce destined to human feed (Franz *et al.*, 2008). Moreover, animal manure has been frequently applied as a fertilizer to land used for crop, increasing the risk

for trasfer of these pathogens to the farm environment, if the manure is not properly matured (Caprioli *et al.*, 2005).

The capability of STEC to survive for long periods, ranging from some weeks to many months, in several environments, such as water, soil and manure (LeJeune *et al.*, 2001, Gagliardi *et al.*, 2002, Caprioli *et al.*, 2005) represents an important issue for horizontal spread of STEC strains and for a recontamination of the same animals.

In the last years the possibility of new environmental transmission pathways emerged. This revolve around the use of biomass originating from wastewater treatment plants or biofuel production systems to improve soils fertility. This biomass concerns the waste originating from urban settements as well as from industrial and zootechnical and agricultural settings (Saveyn *et al.*, 2014), therefore, in principle their use may favour the spreading of STEC and other pathogenic *E. coli* in pastures and farmlands.

Additionally, it has to be considered that free phages carrying the *stx* genes were described as being present in sewage (Muniesa *et al.*, 2004), thus the use of such matter and the related bio-waste in agricoltural settings could increase the occurrence of *stx*-phages in the soil. Here, such phages can act as vectors for the transfer of *stx* genes to different *E. coli* strains, or even other bacteria, leading to the emergence and propagation of new STEC clones and other Stx-producing bacteria. Finally, the spread of both urban and zootechnical waste to the farmlands can lead to the coexistance of human and animal *E. coli* strains, belonging to different pathogroups, in the same environment. This may favour the exchange of genetic material between such *E. coli* populations and consequently the formation of new pathotypes with shuffled virulence genes combination may be favoured. An example of STEC with shuffled virulence features was the STEC O104:H4 strain tha caused in 2011 a large outbreak of infections in Germany and France (Frank *et al.*, 2011). The episode was traced back to the consumption of fenugreek sprouts, and involved more tha 4000 illness cases and more than 800 HUS cases, with the heavy toll of 50 deaths (Frank *et al.*, 2011). The outbreak strain presented a mixed asset of virulence determinats, with a chromosomal backbone of an enteroaggregative *E. coli* strain (EAEC) and genes encoding the Stxs (Frank *et al.*, 2011).

1.6 References

Adachi, J.A., Ericsson, C.D., Jiang, Z.D., DuPont, M.W., Pallegar, S.R. and DuPont, H.L. (2002). Natural history of enteroaggregative and enterotoxigenic Escherichia coli infection among US travelers to Guadalajara, Mexico. *J Infect Dis* **185,** 1681-1683.

Baldini, M.M., Kaper, J.B., Levine, M.M., Candy, D.C. and Moon, H.W. (1983). Plasmid-mediated adhesion in enteropathogenic Escherichia coli. *J Pediatr Gastroenterol Nutr* **2,** 534-538.

Basu, D., Li, X.P., Kahn, J.N., May, K.L., Kahn, P.C. and Tumer, N.E. (2015). The A1 Subunit of Shiga Toxin 2 Has Higher Affinity for Ribosomes and Higher Catalytic Activity than the A1 Subunit of Shiga Toxin 1. *Infect Immun* **84,** 149-161.

Beutin, L., Bode, L., Ozel, M. and Stephan, R. (1990). Enterohemolysin production is associated with a temperate bacteriophage in Escherichia coli serogroup O26 strains. *J Bacteriol* **172,** 6469-6475.

Bien, J., Sokolova, O. and Bozko, P. (2012). Role of Uropathogenic Escherichia coli Virulence Factors in Development of Urinary Tract Infection and Kidney Damage. *Int J Nephrol* **2012,** 681473.

Bray, J. (1945). Isolation of antigenically homogeneous strains of *Bact. coli neapolitanum* from summer diarrhoea of infants*The journal of Pathology* **57,** 239-247.

Brunder, W., Schmidt, H. and Karch, H. (1996). KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic Escherichia coli O157:H7. *Microbiology* **142 (Pt 11),** 3305-3315.

Brunder, W., Schmidt, H. and Karch, H. (1997). EspP, a novel extracellular serine protease of enterohaemorrhagic Escherichia coli O157:H7 cleaves human coagulation factor V. *Mol Microbiol* **24,** 767-778.

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.

Croxen, M.A. and Finlay, B.B. (2010). Molecular mechanisms of Escherichia coli pathogenicity. *Nat Rev Microbiol* **8,** 26-38.

Curtis, M.M., Hu, Z., Klimko, C., Narayanan, S., Deberardinis, R. and Sperandio, V. (2014). The gut commensal Bacteroides thetaiotaomicron exacerbates enteric infection through modification of the metabolic landscape. *Cell Host Microbe* **16,** 759-769.

Del Canto, F. and G., T.A. (2014) Colonization Factor Antigens of Enterotoxigenic Escherichia coli" in Pathogenic Escherichia coli, Molecular and Cellular Microbiology, Caister Academic Press pp. 201-216.

Dell'Omo, G., Morabito, S., Quondam, R., Agrimi, U., Ciuchini, F., Macrì, A. and Caprioli, A. (1998). Feral pigeons as a source of verocytotoxin-producing Escherichia coli. *Vet Rec* **142,** 309-310.

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.

Escher, M., Scavia, G., Morabito, S., Tozzoli, R., Maugliani, A., Cantoni, S.*, et al.* (2014). A severe foodborne outbreak of diarrhoea linked to a canteen in Italy caused by enteroinvasive Escherichia coli, an uncommon agent. *Epidemiol Infect* **142,** 2559-2566.

Escherich, T. (1989). The intestinal bacteria of the neonate and breast-fed infant. 1885. *Rev Infect Dis* **11,** 352-356.

Ethelberg, S., Olsen, K.E., Scheutz, F., Jensen, C., Schiellerup, P., Enberg, J.*, et al.* (2004). Virulence factors for hemolytic uremic syndrome, Denmark. *Emerg Infect Dis* **10,** 842-847.

EUSR (2016). EU summary report on zoonoses, zoonotic agents and food-borne outbreaks 2015 *EFSA Journal* **14**.

Ferreyra, J.A., Wu, K.J., Hryckowian, A.J., Bouley, D.M., Weimer, B.C. and Sonnenburg, J.L. (2014). Gut microbiota-produced succinate promotes C. difficile infection after antibiotic treatment or motility disturbance. *Cell Host Microbe* **16,** 770-777.

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. and van Bruggen, A.H. (2008). Ecology of E. coli O157:H7 and Salmonella enterica in the primary vegetable production chain. *Crit Rev Microbiol* **34,** 143-161.

Friedrich, A.W., Bielaszewska, M., Zhang, W.L., Pulz, M., Kuczius, T., Ammon, A. and Karch, H. (2002). Escherichia coli harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms. *J Infect Dis* **185,** 74-84.

Friesema, I., van der Zwaluw, K., Schuurman, T., Kooistra-Smid, M., Franz, E., van Duynhoven, Y. and van Pelt, W. (2014). Emergence of Escherichia coli encoding Shiga toxin 2f in human Shiga toxin-producing E. coli (STEC) infections in the Netherlands, January 2008 to December 2011. *Euro Surveill* **19,** 26-32.

Friesema, I.H., Keijzer-Veen, M.G., Koppejan, M., Schipper, H.S., van Griethuysen, A.J., Heck, M.E. and van Pelt, W. (2015). Hemolytic uremic syndrome associated with Escherichia coli O8:H19 and Shiga toxin 2f gene. *Emerg Infect Dis* **21,** 168-169.

Frost, L.S., Leplae, R., Summers, A.O. and Toussaint, A. (2005). Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* **3,** 722-732.

Fujiwara, S., Hashiba, H., Hirota, T. and Forstner, J.F. (2001). Inhibition of the binding of enterotoxigenic Escherichia coli Pb176 to human intestinal epithelial cell line HCT-8 by an extracellular protein fraction containing BIF of Bifidobacterium longum SBT2928: suggestive evidence of blocking of the binding receptor gangliotetraosylceramide on the cell surface. *Int J Food Microbiol* **67,** 97-106.

Gagliardi, J.V. and Karns, J.S. (2002). Persistence of Escherichia coli O157:H7 in soil and on plant roots. *Environ Microbiol* **4,** 89-96.

Gamage, S.D., Strasser, J.E., Chalk, C.L. and Weiss, A.A. (2003). Nonpathogenic Escherichia coli can contribute to the production of Shiga toxin. *Infect Immun* **71,** 3107-3115.

Garmendia, J., Frankel, G. and Crepin, V.F. (2005). Enteropathogenic and enterohemorrhagic Escherichia coli infections: translocation, translocation, translocation. *Infect Immun* **73,** 2573-2585.

Grys, T.E., Siegel, M.B., Lathem, W.W. and Welch, R.A. (2005). The StcE protease contributes to intimate adherence of enterohemorrhagic Escherichia coli O157:H7 to host cells. *Infect Immun* **73,** 1295-1303.

Gueimonde, M., Margolles, A., de los Reyes-Gavilán, C.G. and Salminen, S. (2007). Competitive exclusion of enteropathogens from human intestinal mucus by Bifidobacterium strains with acquired resistance to bile--a preliminary study. *Int J Food Microbiol* **113,** 228-232.

Hartland, E.L., Daniell, S.J., Delahay, R.M., Neves, B.C., Wallis, T., Shaw, R.K.*, et al.* (2000). The type III protein translocation system of enteropathogenic Escherichia coli involves EspA-EspB protein interactions. *Mol Microbiol* **35,** 1483-1492.

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.

Jackson, M.P., Newland, J.W., Holmes, R.K. and O'Brien, A.D. (1987). Nucleotide sequence analysis of the structural genes for Shiga-like toxin I encoded by bacteriophage 933J from Escherichia coli. *Microb Pathog* **2,** 147-153.

Jerse, A.E., Yu, J., Tall, B.D. and Kaper, J.B. (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.

Kamada, N., Kim, Y.G., Sham, H.P., Vallance, B.A., Puente, J.L., Martens, E.C. and Núñez, G. (2012). Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. *Science* **336,** 1325-1329.

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

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.

Karmali, M.A., Petric, M., Lim, C., Fleming, P.C., Arbus, G.S. and Lior, H. (1985). The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing Escherichia coli. *J Infect Dis* **151,** 775-782.

Karmali, M.A., Steele, B.T., Petric, M. and Lim, C. (1983). Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing Escherichia coli in stools. *Lancet* **1,** 619-620.

Kassam, Z., Lee, C.H., Yuan, Y. and Hunt, R.H. (2013). Fecal microbiota transplantation for Clostridium difficile infection: systematic review and meta-analysis. *Am J Gastroenterol* **108,** 500-508.

Kenny, B., DeVinney, R., Stein, M., Reinscheid, D.J., Frey, E.A. and Finlay, B.B. (1997). Enteropathogenic E. coli (EPEC) transfers its receptor for intimate adherence into mammalian cells. *Cell* **91,** 511-520.

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.

Konowalchuk, J., Speirs, J.I. and Stavric, S. (1977). Vero response to a cytotoxin of Escherichia coli. *Infect Immun* **18,** 775-779.

Lagier, J.-C. (2016). Gut microbiota and *Clostridium difficile* infections. *Human Microbiome Journal* **2,** 10-14.

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.

LeJeune, J.T., Besser, T.E. and Hancock, D.D. (2001). Cattle water troughs as reservoirs of Escherichia coli O157. *Appl Environ Microbiol* **67,** 3053-3057.

Leyton, D.L., Sloan, J., Hill, R.E., Doughty, S. and Hartland, E.L. (2003). Transfer region of pO113 from enterohemorrhagic Escherichia coli: similarity with R64 and identification of a novel plasmid-encoded autotransporter, EpeA. *Infect Immun* **71,** 6307-6319.

Mansan-Almeida, R., Pereira, A.L. and Giugliano, L.G. (2013). Diffusely adherent Escherichia coli strains isolated from children and adults constitute two different populations. *BMC Microbiol* **13,** 22.

Moon, H.W., Whipp, S.C., Argenzio, R.A., Levine, M.M. and Giannella, R.A. (1983). Attaching and effacing activities of rabbit and human enteropathogenic Escherichia coli in pig and rabbit intestines. *Infect Immun* **41,** 1340-1351.

Morabito, S., Dell'Omo, G., Agrimi, U., Schmidt, H., Karch, H., Cheasty, T. and Caprioli, A. (2001). Detection and characterization of Shiga toxin-producing Escherichia coli in feral pigeons. *Vet Microbiol* **82,** 275-283.

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.

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

Nataro, J.P., Kaper, J.B., Robins-Browne, R., Prado, V., Vial, P. and Levine, M.M. (1987). Patterns of adherence of diarrheagenic Escherichia coli to HEp-2 cells. *Pediatr Infect Dis J* **6,** 829-831.

Nicholson, B.A., West, A.C., Mangiamele, P., Barbieri, N., Wannemuehler, Y., Nolan, L.K*., et al.* (2016). Genetic Characterization of ExPEC-Like Virulence Plasmids among a Subset of NMEC. *PLoS One* **11,** e0147757.

Nielsen, E.M., Skov, M.N., Madsen, J.J., Lodal, J., Jespersen, J.B. and Baggesen, D.L. (2004). Verocytotoxin-producing Escherichia coli in wild birds and rodents in close proximity to farms. *Appl Environ Microbiol* **70,** 6944-6947.

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.

O'Brien, A.D. and Holmes, R.K. (1987). Shiga and Shiga-like toxins. *Microbiol Rev* **51,** 206-220.

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.

O'Brien, A.O., Lively, T.A., Chen, M.E., Rothman, S.W. and Formal, S.B. (1983). Escherichia coli O157:H7 strains associated with haemorrhagic colitis in the United States produce a Shigella dysenteriae 1 (SHIGA) like cytotoxin. *Lancet* **1,** 702.

Ochoa, T.J. and Cleary, T.G. (2003). Epidemiology and spectrum of disease of Escherichia coli O157. *Curr Opin Infect Dis* **16,** 259-263.

Opintan, J.A., Bishar, R.A., Newman, M.J. and Okeke, I.N. (2010). Carriage of diarrhoeagenic Escherichia coli by older children and adults in Accra, Ghana. *Trans R Soc Trop Med Hyg* **104,** 504-506.

Paton, A.W., Srimanote, P., Talbot, U.M., Wang, H. and Paton, J.C. (2004). A new family of potent AB(5) cytotoxins produced by Shiga toxigenic Escherichia coli. *J Exp Med* **200,** 35-46.

Paton, A.W., Srimanote, P., Woodrow, M.C. and Paton, J.C. (2001). Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic Escherichia coli strains that are virulent for humans. *Infect Immun* **69,** 6999-7009.

Persson, S., Olsen, K.E., Ethelberg, S. and Scheutz, F. (2007). Subtyping method for Escherichia coli shiga toxin (verocytotoxin) 2 variants and correlations to clinical manifestations. *J Clin Microbiol* **45,** 2020-2024.

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.

Prager, R., Fruth, A., Siewert, U., Strutz, U. and Tschäpe, H. (2009). Escherichia coli encoding Shiga toxin 2f as an emerging human pathogen. *Int J Med Microbiol* **299,** 343-353.

Pruimboom-Brees, I.M., Morgan, T.W., Ackermann, M.R., Nystrom, E.D., Samuel, J.E., Cornick, N.A. and Moon, H.W. (2000). Cattle lack vascular receptors for Escherichia coli O157:H7 Shiga toxins. *Proc Natl Acad Sci U S A* **97,** 10325-10329.

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.

Rolhion, N. and Chassaing, B. (2016). When pathogenic bacteria meet the intestinal microbiota. *Philos Trans R Soc Lond B Biol Sci* **371**.

Russo, T.A. and Johnson, J.R. (2000). Proposal for a new inclusive designation for extraintestinal pathogenic isolates of Escherichia coli: ExPEC. *J Infect Dis* **181,** 1753-1754.

Sandvig, K. (2001). Shiga toxins. *Toxicon* **39,** 1629-1635.

Sandvig, K., Garred, O., Prydz, K., Kozlov, J.V., Hansen, S.H. and van Deurs, B. (1992). Retrograde transport of endocytosed Shiga toxin to the endoplasmic reticulum. *Nature* **358,** 510-512.

Saveyn, H. and Eder, P. (2014) End-of-waste criteria for biodegradable waste subjected to biological treatment (compost & digestate): Technical proposals. EU Commission, IPTS Sevilla (Spain).

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.

Schmidt, H., Scheef, J., Morabito, S., Caprioli, A., Wieler, L.H. and Karch, H. (2000). A new Shiga toxin 2 variant (Stx2f) from Escherichia coli isolated from pigeons. *Appl Environ Microbiol* **66,** 1205-1208.

Shin, R., Suzuki, M. and Morishita, Y. (2002). Influence of intestinal anaerobes and organic acids on the growth of enterohaemorrhagic Escherichia coli O157:H7. *J Med Microbiol* **51,** 201-206.

Shreiner, A.B., Kao, J.Y. and Young, V.B. (2015). The gut microbiome in health and in disease. *Curr Opin Gastroenterol* **31,** 69-75.

Smith, H.W. and Gyles, C.L. (1970). The relationship between two apparently different enterotoxins produced by enteropathogenic strains of Escherichia coli of porcine origin. *J Med Microbiol* **3,** 387-401.

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

Strockbine, N.A., Jackson, M.P., Sung, L.M., Holmes, R.K. and O'Brien, A.D. (1988). Cloning and sequencing of the genes for Shiga toxin from Shigella dysenteriae type 1. *J Bacteriol* **170,** 1116-1122.

Tatsuno, I., Horie, M., Abe, H., Miki, T., Makino, K., Shinagawa, H*., et al.* (2001). toxB gene on pO157 of enterohemorrhagic Escherichia coli O157:H7 is required for full epithelial cell adherence phenotype. *Infect Immun* **69,** 6660-6669.

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.

Vial, P.A., Robins-Browne, R., Lior, H., Prado, V., Kaper, J.B., Nataro, J.P.*, et al.* (1988). Characterization of enteroadherent-aggregative Escherichia coli, a putative agent of diarrheal disease. *J Infect Dis* **158,** 70-79.

Wallace, J.S., Cheasty, T. and Jones, K. (1997). Isolation of vero cytotoxin-producing Escherichia coli O157 from wild birds. *J Appl Microbiol* **82,** 399-404.

Wiles, T.J., Kulesus, R.R. and Mulvey, M.A. (2008). Origins and virulence mechanisms of uropathogenic Escherichia coli. *Exp Mol Pathol* **85,** 11-19.

Wong, A.R., Pearson, J.S., Bright, M.D., Munera, D., Robinson, K.S., Lee, S.F.*, et al.* (2011). Enteropathogenic and enterohaemorrhagic Escherichia coli: even more subversive elements. *Mol Microbiol* **80,** 1420-1438.

Chapter 2

Aims of the work

2. Aims of the work

STEC are considered a serious public health problem worldwide, because may cause severe diseases, such as HUS, that unfortunately in some cases resolve with the death of human host. As the pathogenesis of STEC infections is not completely understood, the necessity to come to a rapid diagnosis and the comprehension of the following mechanisms of the progression of the disease appear crucial.

The aim of the present work was to investigate STEC populations whose impact on public health was not completely clear, as well as to examine novel or neglected reservoirs and transmission pathways for STEC infections and to investigate the relationships of STEC with the microbial communities, once they establish in the human intestine upon infection. In detail, we studied the possibility that Stx2f-producing STEC could be agent of HUS as well as the prospect that their natural reservoir, pigeons, may be impact public health, being source of HUS-associated strains. Additionally, we adopted holistic state-of-the-art investigative approach, metagenomics, to evaluate the possibility of environmental transmission of the infections through the use of biosolids derived from sludge from wastewater treatment plants and composting plants and used to amend soils for agricultural purposes. Finally, to complete the investigation on the whole axis of human STEC exposure through disease, we used an enhanced metagenomic approach to study the interactions between STEC and the microbiota inhabiting the human intestine upon infection.

The complex symptomatological scenario associated with the intestinal colonization by STEC is mainly due to the huge variability of their virulence genes content, as a consequence of the high genomic plasticity of *E. coli*. The development of high throughput technologies has made it easier the identification of novel pathogenic strains. The Chapter 3 of this work described the use of whole genome sequencing method to characterize strains of Stx2f-producing *E. coli*, isolated from diarrhoea, HUS cases and the natural reservoir. The *E. coli* strains carrying the *stx2f* gene are carried by pigeons, and the accepted paradigm was that these strains may be transmitted to humans, where they generally cause no symptoms or, occasionally, mild diarrhoea. We have demonstrated that the Stx2f-producing STEC can cause HUS under certain circumstances and formulated the hypothesis that pigeons may be a reservoir of highly pathogenic STEC or of their *stx2f*-phages, which may be released in the environment and in turn acquired by *E. coli* strains presenting a virulence genes background favouring the induction of HUS.

This “environmental lead” is interesting and could be a large investigation field when it comes to STEC infections. It is known that the environment represents an important source of pathogenic *E. coli*, given their ability to adapt and survive in a wide spectrum of ecological niches. In Chapters 4 and 5 we investigated if the organic matter produced at wastewater treatment plants (biosolids) or

derived from composting and used in agriculture as soil improvers, may be considered a new way to disseminate pathogenic microorganisms in the environment and therefore a new source of *E. coli*, infections for humans, including those caused by STEC. The environment is a complex ecosystem, inhabited by a huge diversity of microbial species, many of which have not been identified yet. The spread of pathogenic strains of *E. coli* may also alter the microbial equilibrium in the soil and favour the emergence of novel pathogenic microorganisms, representing an important risk for human health. In Chapter 5 a metagenomic approach was developed and used to fully characterize the microbial hazards present in biosolids, including antimicrobial resistance, with emphasis on the presence of STEC and their virulence genes.

The metagenomics is a recent development in the study of complex microbial populations and relies on the direct sequencing of DNA extracted from a given matrix and including nucleic acids from all the species present in the sample.

An enhanced metagenomic approach was employed in Chapter 6 to define the composition of the human intestinal microbiota in presence of STEC infections. Because the pathogenic process of STEC is not completely clear, we investigated whether the infections alters the interactions between STEC and commensal microorganisms in the intestine, eventually influencing the severity of disease. The comprehension of the variations in the microbiota composition between healthy subjects and diseased patients, studied in a set of samples collected during a single outbreak of STEC O26 infections, would be of help in unravelling the roles of commensal bacteria as well as in designing strategies for the possible treatment to re-establish the beneficial species and reduce the clinical manifestations in the patients.

Chapter 3

Whole-Genome Characterization and Strain Comparison of VT2f-Producing *Escherichia coli* Causing Hemolytic Uremic Syndrome

3. Whole-Genome Characterization and Strain Comparison of VT2f- Producing *Escherichia coli* Causing Hemolytic Uremic Syndrome

Shiga toxin-producing *E. coli* are zoonotic pathogens, able to cause a wide spectrum of human diseases, ranging from mild diarrhoea, haemorragic colitis (HC), up to the haemolytic uremic syndrome (HUS). Severe afflictions in humans are due to the STEC ability to produce potent cytotoxins, Shiga toxins (Stx), acquired upon infection with bacteriophages carrying *stx* genes. Several studies indicated that pigeons are the natural reservoir for a sub-group of STEC strains producing the Stx2f variant, which was previously associated with asymptomatic infections or uncomplicated diarhoea in human, hence with a limited impact on public health. Nevertheless, in recent times the number of Stx2f-STEC strains isolated from cases of HUS, appears to have increased.

Whole genome sequencing approach was used in this study to characterize Stx2f-STEC strains isolated from human patients with diarrhoea or HUS and from healthy pigeons. The Virulotyping analysis highlighted the existence of three distinct subpopulations of Stx2f-producing *E. coli*, with isolates from pigeons and from mild diarrhoea specimens displaying a more similar virulence genes content, which do not include all the virulence genes profiles identified in the strains from HUS. In detail, all the Stx2f strains were positive for the *eae* gene, confirming the presence of the LEE-PAI. On the contrary genes implicated in the most severe forms of the infections, such as those conveyed by the OI-122 pathogenicity island (*efa1* gene) and the pO157 virulence plasmid (*ehxA*, *espP* and *katP* genes) were identified only in the isolates from HUS cases.

The results suggested that Stx2f-producing STEC can cause severe human disease, as HUS, in presence of a complete repertoire of virulence-associated determinants. Pigeons should thus be considered reservoir of highly pathogenic STEC strains, or of *stx2f*-phages, which can be released in the environment and in turn lysogenize *E. coli* strains with a virulence genes background associated with HUS, representing a serious risk for public health.

Whole-Genome Characterization and Strain Comparison of VT2f–Producing *Escherichia coli* Causing Hemolytic Uremic Syndrome

Laura Grande, Valeria Michelacci, Roslen Bondì, Federica Gigliucci, Eelco Franz, Mahdi Askari Badouei, Sabine Schlager, Fabio Minelli, Rosangela Tozzoli, Alfredo Caprioli, Stefano Morabito

Verotoxigenic *Escherichia coli* infections in humans cause disease ranging from uncomplicated intestinal illnesses to bloody diarrhea and systemic sequelae, such as hemolytic uremic syndrome (HUS). Previous research indicated that pigeons may be a reservoir for a population of verotoxigenic *E. coli* producing the VT2f variant. We used whole-genome sequencing to characterize a set of VT2f-producing *E. coli* strains from human patients with diarrhea or HUS and from healthy pigeons. We describe a phage conveying the *vtx2f* genes and provide evidence that the strains causing milder diarrheal disease may be transmitted to humans from pigeons. The strains causing HUS could derive from VT2f phage acquisition by *E. coli* strains with a virulence genes asset resembling that of typical HUS-associated verotoxigenic *E. coli*.

Verotoxigenic *Escherichia coli* (VTEC) infections in humans cause a wide spectrum of clinical manifestations ranging from uncomplicated forms of intestinal illnesses to bloody diarrhea and systemic sequelae, such as hemolytic uremic syndrome (HUS) (*1*). The most severe forms are caused by the damage inflicted by the verocytotoxins (VTs) to the target cells in the intestinal mucosa and the renal blood vessels (*1*). The genes encoding the verocytotoxins (*vtx*) are harbored by lambdoid bacteriophages, which can be transferred to multiple bacterial hosts, generating a great diversity in the bacterial types that produce such toxins (*2*).

The most well-known VTEC serogroup, O157, inhabits the gastrointestinal tract of ruminants, especially cattle. However, this and other VTEC serotypes have been isolated from the feces of several other animal species, including deer, pigs, horses, cats, dogs, and wild birds (*3*).

During a program aimed at the control of the pigeon population in Rome, Italy during 1998, G. Dell'Omo et al. observed that this animal species was a carrier of VTEC (*4*). In that study, VTEC of multiple serogroups were isolated from ≈10% of the animals tested. Of 16 VTEC, 15 carried the *eae* gene encoding the intimin and featured genetic determinants that produced a subtype of verocytotoxin type 2 not described before, later designated VT2f (*4–6*). The finding of such a high prevalence of VTEC in pigeons living in Rome led to further research into these bacteria in this and other bird species worldwide. Almost all these studies succeeded in isolating VTEC, with prevalence ranging 3% to >19% in different countries and bird species; most VTEC isolated from pigeon feces and cloacal swab samples harbored the genes encoding the VT2f subtype (*7–10*). These findings emphasize the existence of a strict association between VTEC carrying the *vtx2f* genes and pigeons, which represent a reservoir for such strains.

Data on human illness attributable to VT2f-producing *E. coli* has been scarce until recent reports from Germany and the Netherlands described the isolation of such strains from diarrheal stool specimens from humans (*11*,*12*). Furthermore, in the Netherlands, an HUS case was recently reported to be associated with the presence of a VT2f-producing O8:H19 strain (*13*). We aimed to characterize at the whole-genome level 3 *E. coli* strains that produced the VT2f isolated from HUS and to investigate their relationships with VT2f-producing *E. coli* isolated from human diarrheal cases and from the pigeon reservoir.

Author affiliations: Istituto Superiore di Sanità, Rome, Italy (L. Grande, V. Michelacci, R. Bondi, F. Gigliucci, F. Minelli, R. Tozzoli, A. Caprioli, S. Morabito); National Institute for Public Health and the Environment, Bilthoven, the Netherlands (E. Franz); Islamic Azad University Faculty of Veterinary Medicine, Garmsar Branch, Garmsar, Iran (M. Askari Badouei); Austrian Agency for Health and Food Safety, Graz, Austria (S. Schlager)

DOI: http://dx.doi.org/10.3201/eid2212.160017

Materials and Methods

Bacterial Strains

We investigated 22 Vt2f -producing *E. coli* strains. Eight previously described strains were isolated from pigeons in Italy (*4*); eleven strains were isolated in the Netherlands from fecal specimens from humans with diarrhea during

2008–2012 and are part of the collections held at the National Institute for Public Health and the Environment in the Netherlands (RIVM) (*12*). Of the 3 VT2f-producing *E. coli* from HUS patients, 1 was isolated in Austria in 2013 and 2 in Italy during 2013–2014. A total of 23 unrelated VTEC non-O157 strains that produced VT1 and/or VT2 subtypes other than VT2f have been used for the comparison of the profiles of virulence genes with those of the VT2f-producing isolates (Table 1).

Whole-Genome Sequencing of *E. coli* Strains

Sequencing of the strains isolated from fecal samples from humans with diarrhea and from pigeons was outsourced to the Central Veterinary Institute, Wageningen University (Lelystad, the Netherlands). Genome sequences were obtained by using a TruSeq protocol on an Illumina MiSeq PE300 platform (Illumina, San Diego, CA, USA). The genomes of the 3 VT2f-producing isolates from HUS patients were sequenced by using an Ion Torrent PGM (Thermo Fisher Scientific, Waltham, MA, USA) according to 400-bp protocols for library preparation through enzymatic shearing, Ion OneTouch2 emulsion PCR, enrichment, and Hi-Q sequencing kits (Thermo Fisher Scientific).

The whole-genome sequences (WGSs) of the 23 non-O157 VTEC strains are part of the European Molecular Biology Laboratory's European Nucleotide Archive Study (http://www.ebi.ac.uk/ena; accession no. PRJEB11886). The raw reads have been subjected to quality check through FastQC and trimmed with FASTQ positional and quality trimming tool to remove the adaptors and to accept 20 as the lowest Phred value (*14*).

We subjected the sequences obtained with the Ion Torrent apparatus to de novo assembly by using the tool SPADES (*15*) and those from Illumina by using the A5 pipeline (*16*). The genomes have been assembled in several contigs ranging from 42 to 495 (mean 225), with N50 values (the length of the smallest contig among the set of the largest contigs that together cover at least 50% of the assembly) between 40,736 and 347,638 (mean 152,953). All the contigs were uploaded to the EMBL European Nucleotide Archive (accession no. PRJEB12203). We made annotations by using the Prokka tool (*17*). All the bioinformatics tools used are available on the Aries public Galaxy server (https://w3.iss.it/site/aries/).

Virulence Gene Profile Analysis and Serotyping

The presence of *vtx2f* and *eae* genes has been assessed by PCR by using primers and conditions described elsewhere (*5*,*18*). The activity of VT2f has been evaluated by Vero cell assay (VCA) as previously described (*19*).

We performed detection of the virulence genes *cif, efa1, espABCFIJP, etpD, iha, iss, katP, lpfa, nleABC, tccP, tir, toxB, ehxA*, and *espP* and the serotype determination in silico on the WGSs. We used blastn (available on the Aries public Galaxy server at https://w3.iss.it/site/aries/) to search databases containing the reference sequences of all the known virulence and serotype-associated genes of pathogenic *E. coli* (*20*). To perform the principal component analysis of the virulence gene profiles, we

Table 1. Characteristics of non-O157 verotoxigenic *Escherichia coli* strains used in a comparative analysis of the virulence profile of VT2f-producing strains from humans and the animal reservoir*

Strain	Serogroup	Source†	Year of isolation	Virulence gene profile
ED017	O26	HUS	1989	*eae vtx1*
ED075	O26	Diarrheal feces	1990	*eae vtx1*
ED180	O26	HUS	1994	*eae vtx2*
ED195	O26	HUS	1994	*eae vtx1*
ED392	O26	Diarrheal feces	1998	*eae vtx1*
ED411	O26	HUS	1999	*eae vtx2*
ED423	O26	Diarrheal feces	1999	*eae vtx1*
ED654	O26	HUS	2007	*eae vtx2*
ED669	O26	HUS	2008	*eae vtx1*
ED676	O26	HUS	2008	*eae vtx2*
ED729	O26	Diarrheal feces	2010	*eae vtx1*
ED766	O26	HUS	2010	*eae vtx2*
ED657	O145	HUS	2007	*eae vtx2*
ED603	O121	HUS	2004	*eae vtx2*
ED073	O111	Diarrheal feces	1990	*eae vtx1*
ED082	O111	HUS	1990	*eae vtx1*
ED142	O111	HUS	1992	*eae vtx1 vtx2*
ED178	O111	HUS	1994	*eae vtx1 vtx2*
ED608	O111	HUS	2005	*eae vtx1 vtx2*
ED664	O111	HUS	2007	*eae vtx2*
ED672	O111	HUS	2008	*eae vtx1 vtx2*
ED287	O103	Bovine	1998	*eae vtx1*
ED728	O103	Bloody diarrheal feces	2010	*vtx1*

*All samples are from humans except strain ED287. HUS, hemolytic uremic syndrome.
†HUS samples were isolated from feces.

used SAS/IML studio software version 3.4 (SAS Institute, Inc., Cary, NC, USA).

We investigated plasmid profiles by using PlasmidFinder (*21*; https://cge.cbs.dtu.dk//services/all.php). The intimin subtyping has been performed in silico through a BLAST search (*22*) of the *eae* gene sequences from the WGS against the National Center for Biotechnology Information nucleotide repository. The intimin types of the VT2f-producing strains isolated from pigeons have been published (*6*,*10*).

rpoB Sequencing and Analysis

Amplification and sequencing of the *rpoB* gene were conducted to discriminate between *E. coli* and *E. albertii* species, as previously described (*23*). The amplicons were purified with the SureClean Plus kit (Bioline, London, UK) and sequenced using the BigDye Terminator v1.1 kit on a Genetic Analyzer 3130 (Thermo Fisher Scientific). The obtained sequences were trimmed and aligned to the reference sequences as indicated (*23*), using the Clustal Omega free software (http://www.ebi.ac.uk/Tools/msa/clustalo/).

Typing

We determined *E. coli* phylogenetic groups by using the method of Clermont et al. (*24*). We carried out multilocus sequence typing (MLST) of the VT2f isolates in silico according to the scheme proposed by Wirth et al. (*25*). We analyzed the assembled sequences by using blastn to search the MLST database downloaded from the Internet site of the MLST.UCC Mark Achtman database (http://mlst.warwick.ac.uk/mlst/dbs/Ecoli/).

Single-Nucleotide Polymorphism (SNP) Analysis

We analyzed SNPs by using the tool kSNP3 (*26*) available on the Galaxy project instance Aries (https://w3.iss.it/site/aries/). We set a kmer value of 23.

Results

Characterization of the VT2f-Producing *E. coli* Strains

Serotyping

Of 11 VTEC strains isolated from humans with diarrhea, 5 belonged to the O63:H6 serotype. The remaining 6 isolates contained the $fliC_{H6}$ (3 strains), $fliC_{H7}$ (1 strain), and $fliC_{H34}$ (2 strains) genes (Table 2) and belonged to serogroups O96, O113, O132, O145, and O125. For 1 isolate, the O-antigen–associated genes could not be identified (Table 2) (*12*).

Molecular serotyping of the 8 VT2f-producing strains isolated from pigeons showed that all the isolates had the $fliC_{H2}$ and the O4, O45, O75, and O128 serogroup-associated genes. The O-antigen genes could not be identified for the isolate ED 366 (Table 2). The HUS-associated

Table 2. Characteristics of VT2f-producing *Escherichia coli* investigated in a comparative analysis of the virulence profile of strains isolated from humans with mild and severe disease and from the animal reservoir*

Source and strain	Year isolated	Serotype	Phylotype	MLST	LEE	*adfO*	*efa1*	*cif*	*nleA*	*nleB*	*nleC*	*Hly*	*katP*	*espP*	Intimin type
Human diarrhea															
M856	2008	ONT:H6	B2	ST583	+	+	–	+	–	+	+	–	–	–	α-2
M858	2008	O125:H6	B2	ST583	+	+	–	+	–	+	–	–	–	–	α-2
M859	2009	O113:H6	B2	ST121	+	+	–	+	–	–	–	–	–	–	α-2
M884	2011	O96:H7	B2	ST28	+	+	–	+	+	+	–	–	–	–	β-2
M885	2011	O132:H34	B2	ST582	+	+	–	–	–	+	+	–	–	–	β-2
M900	2012	O145:H34	B2	ST722	+	+	–	–	–	+	–	–	–	–	ι
BCW5711	2012	O63:H6	B2	ST583	+	+	–	+	+	–	+	–	–	–	α-2
BCW5746	2012	O63:H6	B2	ST583	+	+	–	+	–	–	+	–	–	–	α-2
BCW5743	2012	O63:H6	B2	ST583	+	+	–	+	–	–	+	–	–	–	α-2
BCW5739	2012	O63:H6	B2	ST583	+	+	–	+	–	–	+	–	–	–	α-2
BCW5717	2012	O63:H6	B2	ST583	+	+	–	+	–	–	+	–	–	–	α-2
Pigeon															
ED360	1997	O45:H2	B1	ST20	+	+	–	+	+	+	+	–	–	–	β
ED361	1997	O75:H2	B1	ST20	+	+	–	+	+	+	+	–	–	–	β
ED363	1997	O4:H2	B1	UNK	+	+	–	+	+	+	+	–	–	–	β
ED366	1997	ONT:H2	B1	ST2685	+	+	–	+	+	+	+	–	–	–	β
ED369	1997	O45:H2	B1	ST20	+	+	–	+	+	+	+	–	–	–	β
ED377	1997	O4:H2	B1	UNK	+	+	–	+	+	+	+	–	–	–	β
ED430	2000	O45:H2	B1	ST20	+	+	–	+	+	+	+	–	–	–	β
ED444	2000	O128:H2	B1	ST20	+	+	–	+	+	+	+	–	–	–	β
HUS															
EF453	2013	O80:H2	B1	ST301	+	+	+	–	+	+	+	+	–	+	ξ
EF467	2013	O26:H11	B1	ST21	+	+	+	+	+	+	+	+	+	+	β
EF476	2014	O55:H9	B1	ST301	+	+	+	–	+	+	+	+	–	+	ξ

*Human samples were diarrheal or fecal samples from HUS cases and pigeon samples were feces from asymptomatic birds. LEE, locus of enterocyte effacement; MLST, multilocus sequence type; UNK, unknown; +, positive; –, negative.

VT2f-producing *E. coli* strains EF453 and EF476 belonged to serotypes O80:H2 and O55:H9, respectively, while strain EF467 was O26:H11.

Virulence Gene Profiles

The *E. coli* strains carrying the *vtx2f* and isolated from pigeons have been previously reported to produce an active VT2f (*6*). As expected, culture supernatants from VT2f-producing strains isolated from human diarrhea and HUS induced a cytopathic effect on Vero cells morphologically compatible with that caused by verocytoxins.

All the VT2f-producing strains included in the study were positive for the *eae* gene (Table 2) and displayed the presence of the entire locus of enterocyte effacement (LEE) (data not shown). Most of the *E. coli* VT2f-producing strains isolated from diarrheal cases harbored the α-2 intimin type (8/11), followed by the β-2 (2/11) and ι (1/11) types. The VT2f-strains isolated from pigeons had been previously described to have the β-intimin (*6*) in most cases and, more rarely, the α-2 intimin type (*10*). Of 3 HUS-associated VT2f-producing strains, 2 (EF453 and EF476) carried the ξ intimin type and 1 (EF467) had the β intimin (Table 2).

All the pigeon and HUS isolates possessed the complete set of non–LEE-encoded effectors assayed (*nleA, nleB* and *nleC*) (*27*), whereas the isolates from human diarrhea cases displayed an unequal presence of these genes (Table 2). The *efa1* gene, hallmark of the OI-122 pathogenicity island (*28*), was not identified in the isolates from pigeons or from human diarrheal specimens; neither were the genes *ehxA*, *espP* and *katP*, usually present on the large virulence plasmid of VTEC O157 and other VTEC associated with severe human disease (Table 2). However, the gene *adfO*, present on the OI-57 (*29*), was detected in all the strains investigated (Table 2).

The HUS strains EF453, EF467, and EF476 had the entire *efa1* gene. Strain EF467 also had the *ehxA*, *espP*, and *katP* genes; the EF453 and EF476 strains had the *ehxA* and *espP* genes only (Table 2). The analysis of the plasmid profiles substantiated the finding that the 3 HUS-associated strains carried the large virulence plasmid of VTEC, revealing the presence of a sequence 100% homologous to the replicon sequence of the pO26-CRL plasmid from a VTEC O26:H- (GenBank accession no. GQ259888.1), which harbors the genes *ehxA*, *espP*, and *katP*.

On the basis of plasmid profiles analysis, 7 of 11 *E. coli* VT2f-producing strains isolated from human diarrheal feces seemed to have the replicon sequence of the plasmid pSFO (GenBank accession no. AF401292) encoding the enterohemolysin and a cluster of *pap*-like genes called *sfp* in a sorbitol-fermenting *E. coli* O157 (*30*). However, the analysis of the WGSs failed to identify the *ehxA* and the *pap*-like sequences, suggesting that the entire pSFO plasmid was not present.

Principal component analysis of the virulence genes profiles showed that the HUS isolates producing VT2f clustered with the set of non-O157 VTEC isolates used for comparison, rather than with the other VT2f-producing strains (Figure 1). Conversely, the VT2f-producing strains from diarrhea and from pigeons grouped together and apart from the HUS strains (Figure 1).

Phylogenetic Analyses

rpoB Analysis

All the VT2f-producing isolates had an *E. coli*-related *rpoB* sequence (*23*). This finding verified that all the strains investigated were *E. coli*.

Typing

All VT2f-producing *E. coli* isolates from pigeons and the strains isolated from HUS belonged to the B1 phylogenetic group. All the strains isolated from human diarrheal feces were of phylotype B2 (Table 2).

By MLST, most of the pigeon strains investigated (5/8) belonged to sequence type (ST) 20; 1 was ST2685, and 2 were of unknown ST (Table 2), mainly because of the absence of a recognizable *adk* gene sequence. The 5 O63:H6, the 1 O125:H6, and the 1 ONT:H6 VTEC strains from diarrheal fecal specimens belonged to ST583; of the remaining 4 strains, 1 each was of sequence types ST28, ST121, ST582, and ST722 (Table 2).

Of 3 HUS-associated VT2f-producing *E. coli*, 2 (EF453 and EF476) belonged to ST301; strain EF467 was of ST21 (Table 2). All of the STs belonged to different clonal complexes or to any clonal complex, indicating that they were not related each other (data not shown).

SNP Analysis

A parsimony tree representing the core-genome SNPs analysis (Figure 2) shows that VT2f-producing strains from pigeons, human diarrheal feces, and HUS cases cluster apart from each other and from other VTEC strains used for comparison. The HUS-associated EF467 strain clusters together with the group of VTEC non-O157 from human disease, in agreement with the principal component analysis (Figures 1, 2).

Identification of a Bacteriophage Containing the *vtx2f* Genes

The contigs containing the *vtx2f* genes in the different strains ranged 2,500–68,480 bp in size. Upon annotation, they showed the presence of phage-associated genes in the proximity of *vtx2f*, including those encoding the

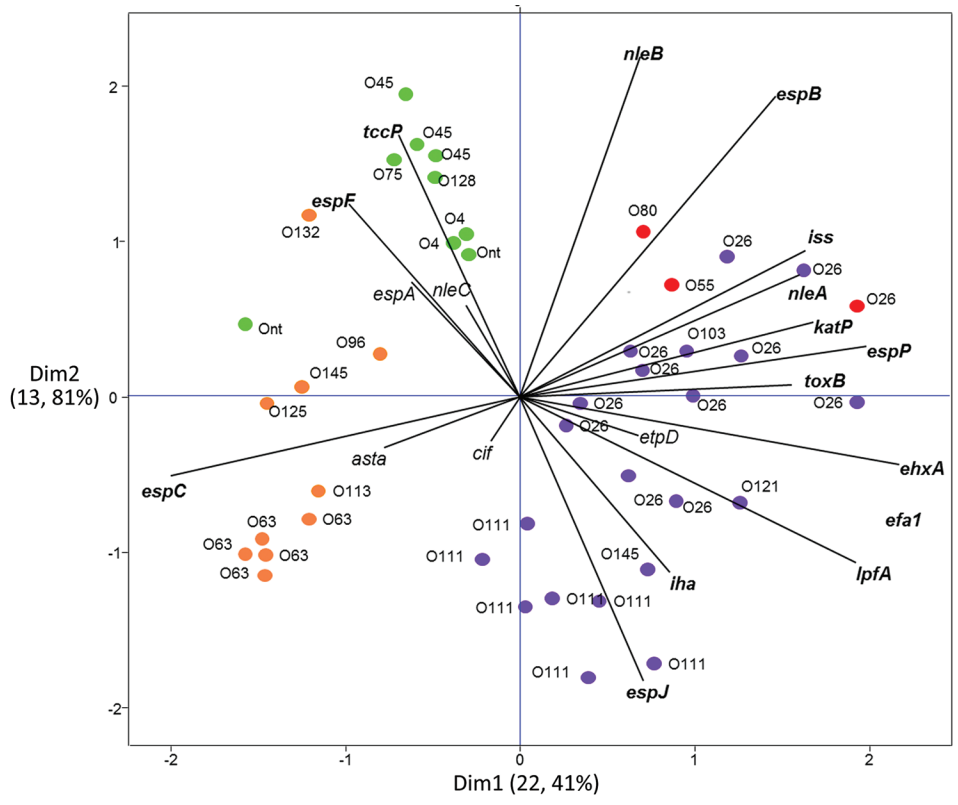


Figure 1. Principal component analysis of the virulence profiles of VT2f-producing verotoxigenic *Escherichia coli* (VTEC) strains isolated from fecal samples from human uncomplicated diarrheal case-patients (orange), human hemolytic uremic syndrome patients (red), and pigeon feces (green). Non-O157 VTEC that do not produce VT2f are indicated in purple. This analysis clusters isolates on the basis of similarities in virulence gene content. Isolates clustering together have similar virulence profiles. Lines represent individual virulence genes. The projected length of the lines on the horizontal and vertical axes indicates the strength of the contribution of a specific virulence marker to the separation of clusters. Dim1 and Dim2 represent the first 2 dimensions of the analysis and indicate the fraction of variation in the dataset that can be explained by the included variables (i.e., virulence genes).

antitermination protein Q, the lysis protein S, a phage terminase, an integrase, and tail-assembly proteins.

We used WGS of HUS strain EF467 to assemble a partial VT2f phage sequence of 38,594 bp (online Technical Appendix, http://wwwnc.cdc.gov/EID/article/22/12/16-0017-Techapp1.pdf). Analysis of the construct highlighted the absence of most of the genes normally involved in the regulation of the switch between the lysogenic state and the lytic cycle of lambdoid phages, such as *cro*, *cI*, *cII*, *cIII*, and N. We confirmed these findings by mapping the raw reads of the WGS of the strain EF467 against the VT2 reference bacteriophage BP933W (GenBank accession no. AF125520) (not shown).

We confirmed the genomic structure of VT2f phage obtained in silico by 4 PCRs performed on the total DNA extracted from strain EF467 by using primer pairs designed on the construct's map and by using a restriction fragment length polymorphism analysis on the obtained PCR fragments (Table 3; online Technical Appendix). The partial VT2f phage sequence was deposited into the EMBL database (accession no. LN997803).

The portion of the VT2f phage spanning the *xerC* gene and the tRNA-Gly and tRNA-Thr loci (online Technical Appendix) was found in the draft genomes of all the VT2f-producing strains. The region downstream to the tRNA loci was also detected, but with different degrees of variation among the WGSs; for example, the presence of an additional DNA stretch of ≈24 kb in the strain BCW5746 (online Technical Appendix).

The presence of a similar phage structure was confirmed by using a long PCR approach in the other 2 VT2f-producing *E. coli* from HUS cases and in 2 strains from pigeons (strains ED377 and ED363). All the strains showed the expected amplicons with PCR 2 and 3 together with the expected product of PCR 1 in 2 pigeon isolates; all the isolates tested did not yield any amplicon with PCR 4 (Figure 3). A comparative analysis of the *xerC* sequences from all the VT2f phage constructs returned a high degree of variation in its sequence, explaining the observed absence of the PCR4-specific amplicon (data not shown). Finally, a BLAST search by using this VT2f phage construct from strain EF467 returned only partial similarity with phages identified in different *Enterobacteriaceae* but did not retrieve similar structures.

Discussion

VTEC producing the VT2f subtype have long been considered a minor public health problem because of their rare association with human infections (*31–34*). Recently, however, an increasing number of reports of human diseases caused by infection with these *E. coli* strains have

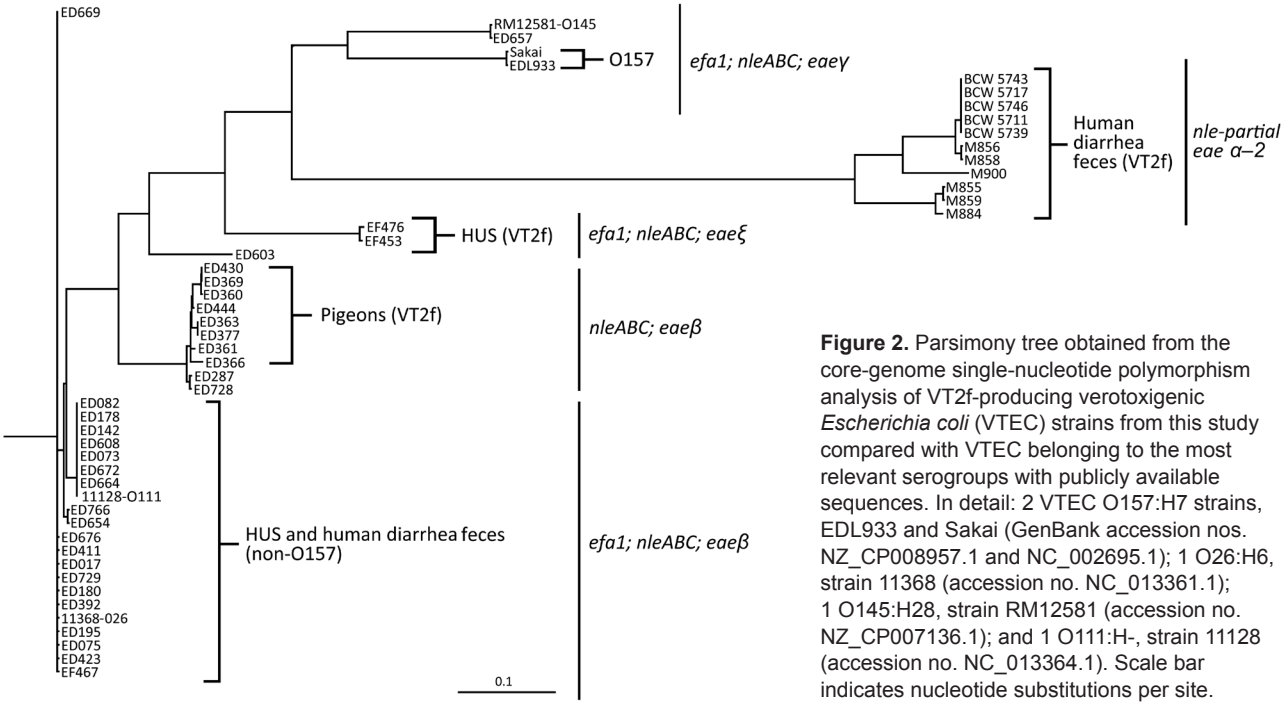


Figure 2. Parsimony tree obtained from the core-genome single-nucleotide polymorphism analysis of VT2f-producing verotoxigenic *Escherichia coli* (VTEC) strains from this study compared with VTEC belonging to the most relevant serogroups with publicly available sequences. In detail: 2 VTEC O157:H7 strains, EDL933 and Sakai (GenBank accession nos. NZ_CP008957.1 and NC_002695.1); 1 O26:H6, strain 11368 (accession no. NC_013361.1); 1 O145:H28, strain RM12581 (accession no. NZ_CP007136.1); and 1 O111:H-, strain 11128 (accession no. NC_013364.1). Scale bar indicates nucleotide substitutions per site.

populated the literature (*11*,*12*,*32*,*35*). The diversity of *vtx2f* gene sequences compared with other *vtx2* subtype genes may have played a role in underestimating the global burden of such infections. PCR primers mostly used for the detection of *vtx2* genes in clinical specimens and the vehicles of infection have been proven to be unable to amplify the *vtx2f* gene (*36*). In addition, the recent description of another *eae*-positive *Escherichia* species often isolated from birds, and sometimes carrying *vtx2f* genes, *E. albertii*, added a further element of confusion. *E. albertii* has been associated both with gastroenteritis in humans and with healthy and diseased birds (*37*,*38*), but this species is difficult to distinguish from *E. coli* when using the usual biochemical or molecular assays.

Most human infections with VTEC producing VT2f have been reported as uncomplicated diarrheal cases (*11*,*12*), which may also have accounted for the underestimation of these infections. Because such cases are not actively surveyed in many countries, these infections may have been overlooked. The recent description of an HUS case associated with a VT2f-producing *E. coli* (*13*) changed the perspective on VT2f-producing *E. coli* and the associated disease, making it necessary to update the current paradigm of HUS-associated VTEC.

Table 3. PCR and restriction fragment length polymorphism analysis conditions used to verify VT2f phage structure in a comparative analysis of the virulence profile of human and zoonotic VT2f-producing *Escherichia coli* strains*

Analysis	Primer name	Sequence, 5′→3′	Position	Thermal profile	Amplicon size, bp	Restriction enzyme (obtained fragments, bp + bp)
PCR1	φ-*vtx2f*_1FW	caccatatcccagcaactgc	1,985–2,005	95°C for 2 min, 30× (94°C for 30 s, 53°C for 30 s, 70°C for 9 min); 72°C for 10 min	6,331	*Pvu*I (1,773 + 4,558)
	φ -*vtx2f*_1RV	gttggcggttccgactacaa	8,315–8,296			
PCR2	φ -*vtx2f*_2FW	gcgcatcaccacttcatctt	8,337–8,357	95°C for 2 min, 30× (94°C for 30 s, 53°C for 30 s, 70°C for 9 min), 72°C for 10 min	8,166	*Hind*III (1,855 + 6,311)
	128–1	agattgggcgtcattcactggttg	16,502–16,479			
PCR3	φ -*vtx2f*_3FW	ggagtggatattgccgacct	16,808–16,827	95°C for 2 min, 30× (94°C for 30 s, 53°C for 30s, 70°C for 9 min), 72°C for 10 min	3,927	*Bgl*II (1,310 + 2,617)
	φ -*vtx2f*_3RV	gtcttcctgctgaggcgatc	20,734–20,715			
PCR4	φ -*vtx2f*_4FW	taatcgcggccgtactcaag	22,172–22,191	95°C for 2 min, 30× (94°C for 30 s, 53°C for 30 s, 70°C for 9 min), 72°C for 10 min	8,808	*Nco*I (5,029 + 3,779)
	φ -*vtx2f*_4RV	tgttcagctccaccttacgg	30,979–30,960			

*Analysis for PCR2, primer 128–1 from (*5*); all other data were compiled for this study. All the long PCR described were performed with the GoTaq Long PCR Master Mix (Promega, Madison, WI, USA) according to manufacturer's instructions. Primer positions refer to the phage sequence deposited into the EMBL database (accession no. LN997803).

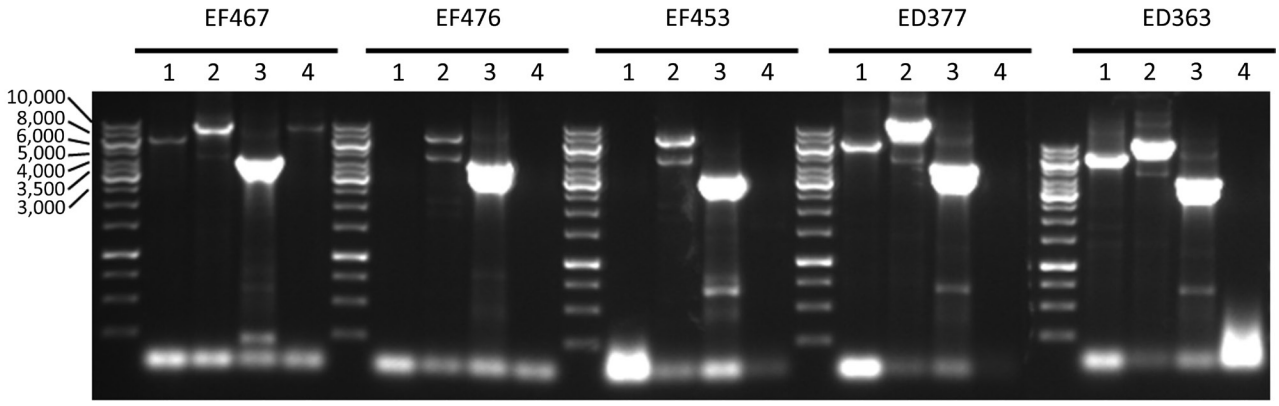


Figure 3. Long PCR analysis of the VT2f phage of verotoxigenic *Escherichia coli* (VTEC) strains isolated from fecal samples from humans with hemolytic uremic syndrome (EF467, EF476, EF453) and from pigeon feces (ED363, ED377). Numbers at left indicate bps; lane numbers indicate PCR 1 to PCR 4. The expected size of the amplicons were 6,331 bp (PCR 1), 8,166 bp (PCR 2), 3,927 bp (PCR 3), and 8,808 bp (PCR 4).

We provide evidence that the VT2f-producing *E. coli* isolated from HUS cases display the complete set of virulence genes described in the typical HUS-associated VTEC (Table 2; Figure 1) (*28*). All VT2f strains from HUS that we examined were positive for pathogenicity island OI-122 (*28*) and the large virulence plasmid first described in VTEC O157 (Table 2) (*39*); the strains from pigeons or from humans with uncomplicated diarrhea did not have these virulence-associated mobile genetic elements (Table 2) (*10–12*,*32*,*40*).

Our study also showed that the LEE was complete in all the genomes investigated, but a complete set of *nleABC* genes was found only in strains from pigeons and from humans with HUS (Table 2), indicating that the VT2f-producing isolates investigated belonged to 3 distinct main virulotypes or subpopulations (Table 2). The intimin subtyping supported this observation. Of 11 diarrheal isolates, 8 had the α-2 gene; all the pigeon isolates had a β intimin coding gene. Finally, 2 of the 3 strains from HUS showed the presence of a gene encoding the ξ intimin (Table 2). Furthermore, the analysis of core genome SNPs confirmed the existence of different subpopulations of VT2f-producing *E. coli* (Figure 2). The analysis of the virulence genes suggests that different populations of VT2f-producing *E. coli* exist and have different potential to cause human disease on the basis of the virulotype to which they belong.

VT2f-producing *E. coli* strains isolated from uncomplicated human cases of diarrhea have been reported in the literature as being ST20 (*11*), which is the same sequence type we identified in most pigeon isolates; this ST was also described in VT2f-producing *E. coli* isolated from pigeons in Japan (*40*). The same study also described an animal isolate of ST722, which was found in 1 strain isolated from human diarrheal feces in our study (Table 2). Similarly, the serotypes in some cases appeared to overlap isolates from pigeons and human cases of diarrhea, such as the serotype O128:H2 that we found in 1 pigeon isolate that was also reported in isolates from human cases of diarrhea in Germany (*11*).

Altogether, these observations indicate that the VT2f-producing *E. coli* causing diarrhea in humans could be a subpopulation of those inhabiting the pigeon reservoir. Alternately, information on the serotypes, ST, and principal component analysis of virulence genes profiles supports the hypothesis that the HUS VT2f-producing strains are more similar to the non-O157 VTEC often isolated from samples from humans with severe disease (Figure 1) than to the other VT2f-producing *E. coli* from humans with diarrhea or from asymptomatic pigeons. This hypothesis suggests that the HUS VT2f-producing strains represent a distinct population of VTEC; whether they are part of the pigeon intestinal flora or arise from an acquisition of the *vtx2*-phage is difficult to ascertain.

The phylogeny of VTEC of different serogroups, investigated by core SNP analysis, showed that the different VT2f-producing *E. coli* cluster into different subpopulations that include strain EF467 grouping together with non-O157 VTEC strains from humans with disease (Figure 2). However, the results from SNP analysis for VTEC of multiple serogroups should be carefully evaluated; the population structure of VTEC belonging to serogroups other than O157 and O26 has not been completely investigated yet.

At the first characterization of the *vtx2f* genes, it was proposed that they were, similar to other VT-coding genes, located on bacteriophages (*5*). Our study confirms this hypothesis and shows that such a phage apparently does not have similar counterparts in the VT-phage genomes reported in the National Center for Biotechnology Information nucleotide repository (http://www.ncbi.nlm.nih.gov/). In addition, we observed that VT2f phage was very similar

in all the VT2f-producing *E. coli* investigated (Figure 3; online Technical Appendix), suggesting that the *vtx2f* genes are present in phages sharing a common ancestor that is different from other phages with the other *vtx1*/*vtx2* subtypes.

In conclusion, we provide evidence that human infections with VT2f-producing *E. coli* are zoonotic diseases transmitted from pigeons. Such an animal reservoir may either directly disseminate VTEC strains causing diarrhea or indirectly release VT2f phages in the environment, which can in turn lysogenize *E. coli* strains that contain accessory virulence determinants and confer them the ability to cause HUS. The isolation of VT2f-producing *E. coli* with a virulence gene profile related to the other HUS-associated VTEC suggests that the severity of the symptoms induced by infection may depend more on the ability to achieve a proficient colonization of the host gut mucosa rather than on the subtype of the produced toxin.

Dr. Grande is a researcher in the field of molecular microbiology. Most of her research activities have been at the European Union Reference Laboratory for *E. coli,* in the unit of Foodborne Zoonoses of the Italian National Institute of Health in Rome, Italy. Her research interests include the investigation and characterization of mobile genetic elements encoding virulence determinants in pathogenic *E. coli*.

References

1. Tozzoli R, Scheutz F. Diarrheagenic *Escherichia coli* infections in umans. In: Morabito S, editor. Pathogenic *Escherichia coli*: molecular and cellular microbiology. Poole (UK): Caister Academic Press; 2014. p. 1–18.
2. Tozzoli R, Grande L, Michelacci V, Ranieri P, Maugliani A, Caprioli A, et al. Shiga toxin–converting phages and the emergence of new pathogenic *Escherichia coli*: a world in motion. Front Cell Infect Microbiol. 2014;4:80. http://dx.doi.org/10.3389/fcimb.2014.00080
3. Caprioli A, Morabito S, Brugere H, Oswald E. Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission. Vet Res. 2005;36:289–311. http://dx.doi.org/10.1051/vetres:2005002
4. Dell'Omo G, Morabito S, Quondam R, Agrimi U, Ciuchini F, Macrì A, et al. Feral pigeons as a source of verocytotoxin-producing *Escherichia coli*. Vet Rec. 1998;142:309–10. http://dx.doi.org/10.1136/vr.142.12.309
5. Schmidt H, Scheef J, Morabito S, Caprioli A, Wieler LH, Karch H. A new Shiga toxin 2 variant (Stx2f) from *Escherichia coli* isolated from pigeons. Appl Environ Microbiol. 2000;66:1205–8. http://dx.doi.org/10.1128/AEM.66.3.1205-1208.2000
6. Morabito S, Dell'Omo G, Agrimi U, Schmidt H, Karch H, Cheasty T, et al. Detection and characterization of Shiga toxin-producing *Escherichia coli* in feral pigeons. Vet Microbiol. 2001;82:275–83. http://dx.doi.org/10.1016/S0378-1135(01)00393-5
7. Nielsen EM, Skov MN, Madsen JJ, Lodal J, Jespersen JB, Baggesen DL. Verocytotoxin-producing *Escherichia coli* in wild birds and rodents in close proximity to farms. Appl Environ Microbiol. 2004;70:6944–7. http://dx.doi.org/10.1128/AEM.70.11.6944-6947.2004
8. Farooq S, Hussain I, Mir MA, Bhat MA, Wani SA. Isolation of atypical enteropathogenic *Escherichia coli* and Shiga toxin 1 and 2f-producing *Escherichia coli* from avian species in India. Lett Appl Microbiol. 2009;48:692–7.
9. Kobayashi H, Kanazaki M, Hata E, Kubo M. Prevalence and characteristics of *eae*- and *stx*-positive strains of *Escherichia coli* from wild birds in the immediate environment of Tokyo Bay. Appl Environ Microbiol. 2009;75:292–5. http://dx.doi.org/10.1128/AEM.01534-08
10. Askari Badouei M, Zahraei Salehi T, Koochakzadeh A, Kalantari A, Tabatabaei S. Molecular characterization, genetic diversity and antibacterial susceptibility of *Escherichia coli* encoding Shiga toxin 2f in domestic pigeons. Lett Appl Microbiol. 2014;59:370–6. http://dx.doi.org/10.1111/lam.12288
11. Prager R, Fruth A, Siewert U, Strutz U, Tschäpe H. *Escherichia coli* encoding Shiga toxin 2f as an emerging human pathogen. Int J Med Microbiol. 2009;299:343–53. http://dx.doi.org/10.1016/j.ijmm.2008.10.008
12. Friesema I, van der Zwaluw K, Schuurman T, Kooistra-Smid M, Franz E, van Duynhoven Y, et al. Emergence of *Escherichia coli* encoding Shiga toxin 2f in human Shiga toxin-producing *E. coli* (STEC) infections in the Netherlands, January 2008 to December 2011. Euro Surveill. 2014;19:26–32. http://dx.doi.org/10.2807/1560-7917.ES2014.19.17.20787
13. Friesema IH, Keijzer-Veen MG, Koppejan M, Schipper HS, van Griethuysen AJ, Heck ME, et al. Hemolytic uremic syndrome associated with *Escherichia coli* O8:H19 and Shiga toxin 2f gene. Emerg Infect Dis. 2015;21:168–9. http://dx.doi.org/10.3201/eid2101.140515
14. Cuccuru G, Orsini M, Pinna A, Sbardellati A, Soranzo N, Travaglione A, et al. Orione, a web-based framework for NGS analysis in microbiology. Bioinformatics. 2014;30:1928–9. http://dx.doi.org/10.1093/bioinformatics/btu135
15. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77. http://dx.doi.org/10.1089/cmb.2012.0021
16. Tritt A, Eisen JA, Facciotti MT, Darling AE. An integrated pipeline for de novo assembly of microbial genomes. PLoS ONE. 2012;7:e42304. http://dx.doi.org/10.1371/journal.pone.0042304
17. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9. http://dx.doi.org/10.1093/bioinformatics/btu153
18. Paton AW, Paton JC. Detection and characterization of Shiga toxigenic *Escherichia coli* by using multiplex PCR assays for stx1, stx2, eaeA, enterohemorrhagic E. coli hlyA, rfbO111, and rfbO157. J Clin Microbiol. 1998;36:598–602.
19. Caprioli A, Nigrelli A, Gatti R, Zavanella M. Isolation of verotoxin-producing *Escherichia coli* from slaughtered pigs. Eur J Clin Microbiol Infect Dis. 1993;12:227–8. http://dx.doi.org/10.1007/BF01967122
20. Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM, Scheutz F. Rapid and easy in silico serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. J Clin Microbiol. 2015;53:2410–26. http://dx.doi.org/10.1128/JCM.00008-15
21. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014;58:3895–903. http://dx.doi.org/10.1128/AAC.02412-14
22. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10. http://dx.doi.org/10.1016/S0022-2836(05)80360-2
23. Lindsey RL, Fedorka-Cray PJ, Abley M, Turpin JB, Meinersmann RJ. Evaluating the occurrence of *Escherichia albertii* in chicken carcass rinses by PCR, Vitek analysis, and sequencing of the *rpo*B gene. Appl Environ Microbiol. 2015;81:1727–34. http://dx.doi.org/10.1128/AEM.03681-14

24. Clermont O, Christenson JK, Denamur E, Gordon DM. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environ Microbiol Rep. 2013;5:58–65. http://dx.doi.org/10.1111/1758-2229.12019
25. Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, et al. Sex and virulence in *Escherichia coli*: an evolutionary perspective. Mol Microbiol. 2006;60:1136–51. http://dx.doi.org/10.1111/j.1365-2958.2006.05172.x
26. Gardner SN, Slezak T, Hall BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. Bioinformatics. 2015;31:2877–8. http://dx.doi.org/10.1093/bioinformatics/btv271
27. Dean P, Kenny B. The effector repertoire of enteropathogenic *E. coli*: ganging up on the host cell. Curr Opin Microbiol. 2009;12:101–9. http://dx.doi.org/10.1016/j.mib.2008.11.006
28. Karmali MA, Mascarenhas M, Shen S, Ziebell K, Johnson S, Reid-Smith R, et al. 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. 2003;41:4930–40. http://dx.doi.org/10.1128/JCM.41.11.4930-4940.2003
29. Imamovic L, Tozzoli R, Michelacci V, Minelli F, Marziano ML, Caprioli A, et al. 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. 2010;78:4697–704. http://dx.doi.org/10.1128/IAI.00512-10
30. Brunder W, Karch H, Schmidt H. Complete sequence of the large virulence plasmid pSFO157 of the sorbitol-fermenting enterohemorrhagic *Escherichia coli* O157:H- strain 3072/96. Int J Med Microbiol. 2006;296:467–74. http://dx.doi.org/10.1016/j.ijmm.2006.05.005
31. Friedrich AW, Bielaszewska M, Zhang WL, Pulz M, Kuczius T, Ammon A, et al. *Escherichia coli* harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms. J Infect Dis. 2002;185:74–84. http://dx.doi.org/10.1086/338115
32. Jenkins C, Willshaw GA, Evans J, Cheasty T, Chart H, Shaw DJ, et al. Subtyping of virulence genes in verocytotoxin-producing *Escherichia coli* (VTEC) other than serogroup O157 associated with disease in the United Kingdom. J Med Microbiol. 2003; 52:941–7. http://dx.doi.org/10.1099/jmm.0.05160-0
33. Seto K, Taguchi M, Kobayashi K, Kozaki S. Biochemical and molecular characterization of minor serogroups of Shiga toxin–producing *Escherichia coli* isolated from humans in Osaka prefecture. J Vet Med Sci. 2007;69:1215–22. http://dx.doi.org/10.1292/jvms.69.1215
34. van Duynhoven YT, Friesema IH, Schuurman T, Roovers A, van Zwet AA, Sabbe LJ, et al. Prevalence, characterisation and clinical profiles of Shiga toxin–producing *Escherichia coli* in the Netherlands. Clin Microbiol Infect. 2008;14:437–45. http://dx.doi.org/10.1111/j.1469-0691.2008.01963.x
35. Buvens G, De Rauw K, Roisin S, Vanfraechem G, Denis O, Jacobs F, et al. Verocytotoxin-producing *Escherichia coli* O128ab:H2 bacteremia in a 27-year-old male with hemolytic-uremic syndrome. J Clin Microbiol. 2013;51:1633–5. http://dx.doi.org/10.1128/JCM.03025-12
36. Feng PC, Jinneman K, Scheutz F, Monday SR. Specificity of PCR and serological assays in the detection of *Escherichia coli* Shiga toxin subtypes. Appl Environ Microbiol. 2011;77:6699–702. http://dx.doi.org/10.1128/AEM.00370-11
37. Oh JY, Kang MS, Hwang HT, An BK, Kwon JH, Kwon YK. Epidemiological investigation of *eae*A-positive *Escherichia coli* and *Escherichia albertii* strains isolated from healthy wild birds. J Microbiol. 2011;49:747–52. http://dx.doi.org/10.1007/s12275-011-1133-y
38. Ooka T, Tokuoka E, Furukawa M, Nagamura T, Ogura Y, Arisawa K, et al. Human gastroenteritis outbreak associated with *Escherichia albertii*, Japan. Emerg Infect Dis. 2013;19:144–6. http://dx.doi.org/10.3201/eid1901.120646
39. Karch H, Heesemann J, Laufs R, O'Brien AD, Tacket CO, Levine MM. A plasmid of enterohemorrhagic *Escherichia coli* O157:H7 is required for expression of a new fimbrial antigen and for adhesion to epithelial cells. Infect Immun. 1987;55:455–61.
40. Murakami K, Etoh Y, Ichihara S, Maeda E, Takenaka S, Horikawa K, et al. Isolation and characteristics of Shiga toxin 2f-producing *Escherichia coli* among pigeons in Kyushu, Japan. PLoS ONE. 2014;9:e86076. http://dx.doi.org/10.1371/journal.pone.0086076

Address for correspondence: Laura Grande, European Union Reference Laboratory for *E. coli*, Veterinary Public Health and Food Safety Department, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome, Italy; email: laur.grande@gmail.com

Chapter 4

Pathogenic *Escherichia coli* and enteric viruses in biosolids and related top soil improvers in Italy

4. Pathogenic *Escherichia coli* and enteric viruses in biosolids and related top soil improvers in Italy

Shiga toxin-producing *E. coli* are zoonotic pathogens and their transmission to humans occurs, mainly, via foodborne and waterborne routes, through inter-human cycles or direct contacts with STEC positive animals, or the environment. Faecal shedding of STEC from asymptomatic or symptomatic animals involves a wide contamination of farmland, water and consequently fresh produce destined to human feed. It is conceivable that the environment represents an important source of pathogenic *E. coli*, given their ability to adapt and survive in a wide spectrum of ecological niches.

Nowadays, the need, to high crop with a reduced usage of water and expensive fertilizers to amend soils, led to the use of Biosolids (BSO), such as biomass resulting from sewage sludge derived from treatment of industrial, municipal and animal farming wastewater as well as compost from green wastes, as soil improvers. Data available on the characterisation of the hazards possibly present in the BSO are scarce, thus the concern related to potential risks for food safety and human health is unknown.

In the present paper, we evaluated the possible presence of zoonotic agents in a set of BSO samples, performing a molecular screening based on Real Time PCR. The analysis gave positive results for virulence genes of human adenovirus, human norovirus and diarrheagenic *E. coli*. Most of the analysed BSO showed positive signals for *stx* genes. This could increase the occurrence of STEC and/or *stx*-phages in the soil, which in turn could be acquired by resident *E. coli* strains, or even by other bacteria, leading to the emergence and propagation of novel STEC clones. Given the mixed nature of the concerned matter, the contemporaneous presence of different strains and phages could facilitate the emergence of pathogenic bacteria, including STEC with shuffled virulence genes combination. As a matter of fact, sewage sludges are sometimes produced in wastewater treatment plants collecting wastes from urban settlements and from livestock. Moreover, according to the regulatory framework applicable, compost can be added with sludges up to 30% in weight.

Additionally, our results suggested the presence in the enrichment culture from BSO, of viable microbial cells of pathogenic *E. coli,* indicating that such soil improvers may a potential source of STEC and other pathogenic *E. coli* and demanding a thorough investigation of the related risks for human health.

Journal of Applied Microbiology

Journal of Applied Microbiology ISSN 1364-5072

ORIGINAL ARTICLE

Pathogenic *Escherichia coli* and enteric viruses in biosolids and related top soil improvers in Italy

R. Tozzoli, I. Di Bartolo, F. Gigliucci, G. Brambilla, M. Monini, E. Vignolo, A. Caprioli and S. Morabito

Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare, Istituto Superiore di Sanità, Rome, Italy

Keywords
Escherichia coli (all potentially pathogenic types), environmental, microbial contamination, sludges, viruses.

Correspondence
Rosangela Tozzoli, Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare, Istituto Superiore di Sanità, Viale Regina Elena 299 00161, Rome, Italy.
E-mail: rosangela.tozzoli@iss.it

2016/1704: received 25 May 2016, revised 3 August 2016 and accepted 22 September 2016

doi:10.1111/jam.13308

Abstract

Aims: To investigate the presence of genomic traits associated with a set of enteric viruses as well as pathogenic *Escherichia coli* in top soil improvers (TSI) from Italy.
Methods and Results: Twenty-four TSI samples originating from municipal sewage sludges, pig manure, green and household wastes were analysed by real time PCR for the presence of hepatitis E virus (HEV), porcine and human adenovirus (HuAdV), norovirus, rotavirus and diarrhoeagenic *E. coli*. None of the samples was found positive for HEV or rotavirus. Four samples were positive for the presence of nucleic acids from human norovirus, two of them being also positive for HuAdV. Real time PCR screening gave positive results for many of the virulence genes characteristic of diarrhoeagenic *E. coli* in 21 samples. These included the verocytotoxin-coding genes, in some cases associated with intimin-coding gene, and markers of enteroaggregative, enterotoxigenic and enteroinvasive *E. coli*.
Conclusions: These results provide evidence that enteric viruses and pathogenic *E. coli* may be released into the environment through the use of sludge-derived TSI.
Significance and Impact of the Study: The results highlight that the TSI-related environmental risk for the food chain should be more deeply assessed.

Introduction

The use of biosolids (BSO) from municipal sewage sludge (MSS), their derived anaerobic digestates (DIG) and biosolids-based composts (MCO) as top soil improvers (TSI) in agriculture, is currently considered of impact within the implementation of the sustainability policies for the reduction of the carbon- and water-footprints in food production systems. The main benefits rely on their potential use as soil conditioners and fertilizers, with organic carbon input supporting high crop and fodder yields per hectare and on a reduced water usage for amended soil (Usman *et al.* 2012). On the other hand, concerns have been raised for the potential risks for food safety and human health. As a matter of fact the possible transfer of microbial and chemical hazards from TSI to soils intended for forage and food production has been suggested (Saveyn and Eder 2014; Rhodes *et al.* 2015). These include microbial pathogens of animal/human origin, including viruses, bacteria and parasites, as well as contaminants of anthropogenic origin, such as pharmaceuticals, heavy metals and persistent organic pollutants from consumer products (EFSA 2014; Saveyn and Eder 2014; Rhodes *et al.* 2015).

In the United States and Canada, about 60% of BSO meeting the regulatory requirements are recycled through agricultural land application (Cogger *et al.* 2006). Within the European Union, it has been estimated that on average 4 Mt year^{-1} MSS of a total 10 Mt year^{-1} production are disposed in agriculture both directly or indirectly, after being subjected to sanitization and stabilization treatment processes, mainly based on biological fermentation (Saveyn and Eder 2014). In the case of DIG, the main treatment includes an anaerobic digestion at temperatures suitable for mesophilic or thermophilic bacteria fermentation, allowing for the production of methane bio-fuel. In the case of MCO, the controlled decomposition of different biodegradable materials coming from

green and household bio-wastes and from their DIG, is performed under predominantly aerobic conditions operated by thermophilic bacteria (Saveyn and Eder 2014) (Fig. 1).

End-of-waste criteria for the safe public use of BSO and derived DIG and MCO have been enforced in some EU Member States on a national basis, but they still lack harmonization. In several EU countries, such as Germany, United Kingdom, the Netherlands and Denmark, the sewage sludge ordinances restrict MSS to few applications, in order to reduce potential health risks (NRC 2002; Wiechman *et al.* 2015). In Italy, BSO basically can be used to dress soils not intended for horticulture, and, in the case of their use, harvesting and grazing is withdrawn for 35 days after the application of BSO. Moreover, the used sludge must fulfil microbiological criteria that include the absence of *Salmonella* in 25 g dry weight (dw) test portions ($n = 5$; $c = 0$) and an *Escherichia coli* load comprising between 1000 and 5000 colony forming units per gram dw of test portion ($n = 5$; $c = 1$) (Lgs 2010). The indirect use of BSO in agriculture is allowed according to the national law (Lgs 2013): MCO should be labelled differently from other green and/or household waste derived MCO; BSO can be present up to 35% (dry matter basis) in the MCO composition, with a maximum allowed use of 35 ton ha^{-1} $year^{-1}$ under a microbial end-of-waste criteria of $<10^3$ MPN per g dw for *Salmonella* and their use is forbidden in organic farms. A national legislative rule for the use of BSO-derived DIG is currently under definition.

The benefit proposed for the use of these materials as TSI has not been counterbalanced by a thorough risk evaluation yet, and the data available on the characterisation of the hazards possibly present in the TSI are scant, at least in the EU, because not all are framed within routine monitoring plans (Saveyn and Eder 2014). The evaluation of the quality of BSO, based on the assessment of the presence of a few indicator micro-organisms, may underestimate the risk of infections (Viau *et al.* 2011). As a matter of fact, TSI may comprise inputs from both urban settlements and intensive farming systems, thus representing a possible vehicle for the transfer of new and emerging pathogens, including agents of zoonoses, to the soils. This has been acknowledged as a potential new epidemiological route for human infections (EFSA 2014). The aim of this study was the characterization of a set of samples of BSO, DIG and MCO for the presence of diarrhoeagenic *E. coli* strains and enteric viruses as markers of the presence of human and zoonotic pathogens.

Materials and methods

Sampling procedure

The TSI samples used in this study ($n = 24$) were collected in 2013 and were composed of BSO ($n = 11$)

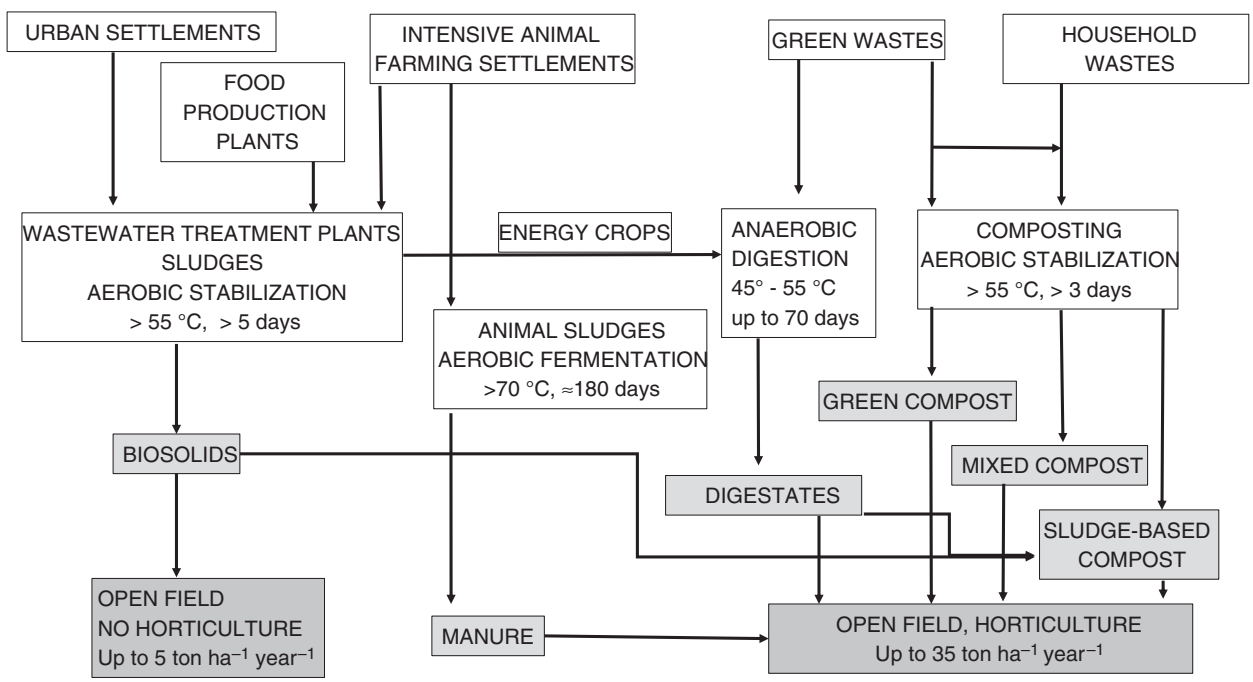


Figure 1 Schematic diagram of the production of biosolids, digestates and composts intended for agriculture soil dressing starting from different sources: urban, food industry and intensive animal farming sludges from wastewater treatment plants, green and household wastes.

derived from different Emilia-Romagna Region tertiary wastewater treatment plants (>25 000 population equivalent, p.e.), with merged inputs from urban and intensively farmed pig settlements; urban BSO recovered from Tuscany ($n = 1$) (>150 000 p.e.), Veneto ($n = 2$) and Piedmont ($n = 1$) regions, respectively (>25 000 p.e.); one sample of manure and the derived DIG obtained from the Sardinia region from a bio-gas plant annexed to an intensive pig farm (>2500 p.e.); MCO, accounting for different contributions from household wastes, green wastes and urban sewage ($n = 5$) from different Italian regions (Fig. 1) and two compost samples made up of green wastes only (GCO) intended for organic horticulture. All the BSO and DIG samples were sampled in accordance with the UNI 10802/2004 guideline and derived from internal/official checks carried out under the provision of the national law for the analysis of heavy metals; MCO and GCO samples were available for purchase on the market. Pig manure and DIG were kindly provided by the farm owner.

BSO from wastewater treatment plants were stored in glass jars, in the dark, at −30°C, until analysis.

Nucleic acid extraction and real time PCR screening for virus detection

Samples were suspended in 20% diethyl-pyrocarbonate water (g per vol) and centrifuged at 10 500 ***g*** for 10 min at 4°C. A murine norovirus strain (MNV-1) was used as control for the nucleic acid extraction process. Five microlitre of a suspension of MNV-1 ($4{\cdot}7 \times 10^7$ PFU per ml) was added to 900 μl of the supernatant and the samples were immediately processed by a Qiamp-viral RNA mini kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. Both DNA and RNA were successfully purified with this approach, as reported previously (Di Bartolo *et al.* 2012). Nucleic acid extracts were assayed immediately or stored at −80°C.

The extracted RNA was subjected to real time reverse transcription PCR (real time RT-PCR) for the detection of MNV-1 as the process control. Samples positive for MNV-1 were further tested by real-time PCR for the presence of human and animal enteric viruses including human norovirus (HuNoV, genogroups I and II), hepatitis E virus (HEV), porcine and human adenoviruses (pAdV and HuAdV) and group A rotavirus (RVA). All reaction mixtures included an internal amplification control (IAC) (Diez-Valcarce *et al.* 2011) with the exception of the RVA real time RT-PCR for which no IAC was available. Each sample was analysed in duplicate. The nucleic acid was used undiluted or following a 10-fold dilution in case the IAC failed to be detected. Details of the single screening assays are given below.

Real time RT-PCR for the detection of MNV-1, HEV and HuNoV

Five microlitre of RNA was analysed for the detection of MNV-1, HEV and HuNoV (both GI and GII) by a duplex one-step reverse-transcription real time PCR detecting both IAC and target virus (Kageyama *et al.* 2003; Jothikumar *et al.* 2006; da Silva *et al.* 2007; Baert *et al.* 2008).

Real time PCR for pAdV and HuAdV

The screening of the samples for the presence of pAdV and HuAdV was carried out by duplex real time PCR as described previously (Hundesa *et al.* 2006). Nucleic acids extracted from samples positive in real time PCR for HuAdV, were subjected for confirmation to conventional PCR followed by nested-PCR to amplify a 171 bp conserved fragment of the hexon gene as described previously (Allard *et al.* 2001).

Real time RT-PCR for rotavirus detection

A real time RT-PCR on the rotavirus NSP3-coding gene was also performed, following the protocol described by Jothikumar *et al.* (2009).

To assess the possible presence of inhibitors of real time RT-PCR, the samples were also analysed after spiking with serial dilutions of RNA extracted from human Wa rotavirus (10^4 PFU).

Human adenovirus replication on A549 cells

A549 (ATCC® CCL-185™) cells were grown in T12 cm^2 flasks in Dulbecco's modified Eagle's medium (D-MEM) supplemented with penicillin (250 U ml^{-1}), streptomycin (250 mg ml^{-1}), 1% glutamine and 10% heat-inactivated fetal calf serum (D-MEM 10% FBS). Three-hundred microlitres of samples 13-0484 and 13-0488 undiluted and a 10-fold dilution were added to the cell monolayers. Mock infected cells were used as a negative control. After 6 h, the inoculum was removed and the flasks were incubated at 37°C and 5% CO_2. The cytopathic effect was checked for 7 days and the daily collected supernatants were subjected to real-time PCR as described above.

Real time PCR screening for the virulence genes of diarrhoeagenic *Escherichia coli*

Real time PCR amplification of genes associated with diarrhoeagenic *E. coli* has been carried out on DNA extracted directly from each sample as well as from the cultures obtained after enrichment in non-selective

medium. The enrichment of the test samples was performed by adding 1 g to 100 ml of buffered peptone water (Biomerieux, Craponne, France) followed by incubation for 18–24 h at 37°C. The DNA was extracted from 1 ml of the enrichment cultures using the InstaGene DNA matrix (Bio-Rad Laboratories, Hercules, CA) and from 0·25 g of untreated TSI samples, by using the PowerSoil® DNA isolation kit (MoBio Laboratories, Carlsbad, CA). Real time PCR amplifications were carried out for detecting the main virulence genes characterizing the different pathogenic *E. coli* types using 2 μl of a 10-fold dilution of the DNA obtained from enrichment cultures and 2 μl of undiluted DNA directly extracted from TSI samples. Samples were analysed in duplicate in each real time PCR assay, and positive controls were included in each analytical session. In particular the verocytotoxin-coding-genes *vtx1* and *vtx2* were detected for the identification of verocytotoxin producing *E. coli* (VTEC) by using the primers and probes of Perelle *et al.* (2004); the *eae* gene encoding the adhesin responsible for colonization shared by enteropathogenic *E. coli* (EPEC) and certain VTEC strains, with the primers described by Nielsen and Andersen (2003); *aggR* and *aaiC* genes were used as markers for the presence of enteroaggregative *E. coli* (EAEC) (primers and probes developed by EU-RL *E. coli*, http://www.iss.it/binary/vtec/cont/EU_RL_VTEC_Method_05_Rev_1.pdf); *ipaH*, a gene located on a large virulence plasmid was amplified to detect enteroinvasive *E. coli* (EIEC); and the *sth*, *stp* and *lt* genes coding for heat-stable and heat-labile enterotoxins were used for the identification of enterotoxigenic *E. coli* (ETEC). The primers and probes for EIEC and ETEC virulence genes detection have been described elsewhere (Wang *et al.* 2010; Liu *et al.* 2013). The IAC described by Fricker *et al.* (2007) was used in all the PCR tests (100 copies per reaction).

Isolation of VTEC from TSI samples

Ten microlitres of the enrichment cultures was plated onto MacConkey and TBX agar plates. From the plates, colonies displaying an *E. coli* morphology were picked up and subjected to real time PCR for the presence of *vtx* genes (Perelle *et al.* 2004).

Results

Screening of biosolids and GCO for the presence of enteric pathogenic viruses

The results of the detection of HEV, HuNoV and RVA nucleic acid in BSO, DIG, GCO and MCO samples is reported in Table 1. Some of the BSO were derived from municipal sewage treatment plants, which in some cases collected inputs from urban and intensively farmed pig settlements. Therefore we have also evaluated the presence of human and swine enteric virus, such as the ubiquitous HuAdV and pAdV. As a matter of fact, the presence of pAdV can be considered a relevant indicator of swine faecal contamination, being normally detected between 90 and 98% of swine faecal samples (Maluquer de Motes *et al.* 2004; Di Bartolo *et al.* 2012).

All samples were positive for the process control (MNV-1) and showed comparable C_t values, indicating that the nucleic acid extraction step was carried out successfully. For two samples it was necessary to analyse a 10-fold dilution of the nucleic acid, since the undiluted sample was inhibited when assayed for the presence of MNV-1 RNA.

Four of the 24 samples analysed, three BSO and one MCO, were positive for HuNoV, showing C_t values ≥36. Two BSO samples contained HuNoV of genogroup GI (13-0495; 13-0476) sequences and two, one BSO and one MCO, were positive for GII (13-0484; 13-0488) (Table 1). The latter two samples were also positive for HuAdV (13-0484; 13-0488). Two additional samples showed inhibition in the real time PCR for HuAdV detection as proven by the lack of curves related to IAC when analysed as undiluted. Sequencing of the amplicon from one of the two positive samples, confirmed the presence of a genome fragment of HuAdV 41 (data not shown). However, the viability of the adenovirus could not be demonstrated as we did not observe replication in A549 cells. Finally, none of the samples were found positive for HEV, pAdV and RVA.

Screening of biosolids and GCO for the presence of diarrhoeagenic *Escherichia coli*-associated genes

Real time PCR detection of the main virulence genes of diarrhoeagenic *E. coli* gave positive results in several samples (Table 1), even though in many cases the signals arose very late (i.e. after 30 C_t). As for the analysis of the DNA extracted directly from the specimens, nine samples tested positive for the presence of *vtx2* gene, together with the *eae* gene in six samples. One additional specimen was positive for the *eae* gene only. Two of 24 samples proved positive for the EIEC gene *ipaH*, and four showed the presence of the *lt* gene coding for the heat labile toxin of ETEC. Finally, nine samples were positive for *aggR* gene of EAEC (Table 1).

Nine of the samples assayed after enrichment, spanning all the different sample types, were positive for the presence of *vtx* genes, characteristic of the VTEC pathogroup. In eight samples the presence of *eae* gene, encoding the intimin adhesin of EPEC in certain VTEC, was also

Table 1 Detection of human adenovirus, human norovirus genomes and virulence genes associated with diarrhoeagenic *Escherichia coli* in nucleic acids extracted from top soil improvers

		Viral genomes			Diarrhoeagenic *E. coli* virulence genes									Diarrhoeagenic *E. coli* virulence genes								
			HuNoV		Nucleic acid extracted from enrichment cultures									Nucleic acid extraction directly from samples								
Sample	Type	HuAdV	GI*	GII	*vtx1*	*vtx2*	*eae*	*aggR*	*aaiC*	*lt*	st_h	st_p	*ipaH*	*vtx1*	*vtx2*	*eae*	*aggR*	*aaiC*	*lt*	st_h	st_p	*ipaH*
13-0474	BSO	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
13-0475	BSO	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
13-0476	BSO	–	+	–	–	+	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–
13-0477	DIG	–	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	–	–	–	–	–
13-0478	BSO	–	–	–	–	+	–	–	–	–	–	–	–	–	+	+	–	–	–	–	–	–
13-0479	BSO	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–
13-0482	MCO	–	–	–	+	+	+	–	–	–	–	–	–	–	+	+	+	–	–	–	–	–
13-0483	MCO	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
13-0484	MCO	+	–	+	–	+	+	–	–	–	–	–	–	–	+	–	+	–	–	–	–	+
13-0485	GCO	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
13-0486	MCO	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–
13-0487	GCO	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
13-0488	BSO	+	–	+	–	–	+	–	–	–	–	–	–	–	+	+	+	–	–	–	–	–
13-0489	BSO	–	–	–	–	+	+	–	–	–	–	–	–	–	+	–	+	–	–	–	–	–
13-0490	BSO	–	–	–	–	–	+	–	–	–	–	–	–	–	+	+	+	–	–	–	–	–
13-0492	BSO	–	–	–	–	+	+	–	–	–	–	–	–	–	+	+	+	–	–	–	–	+
13-0496b	BSO	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
13-0497	MAN	–	–	–	–	+	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–
13-0480	BSO	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–
13-0481	MCO	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–
13-0493	BSO	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–
13-0494	BSO	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–
13-0495	BSO	–	+	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	+	–	–	–
13-0496	BSO	–	–	–	–	–	+	–	–	–	–	–	+	–	–	–	–	–	+	–	–	–

BSO, biosolids; DIG, digestates; GCO, green wastes only; MAN, manure; MCO, biosolid-based compost.

demonstrated. Four of the eight *eae*-positive samples showed the concomitant presence of the *vtx*-coding genes. One sample was positive for the *aggR* gene and *vtx2*, and three showed the presence of *ipaH*, one of which was also positive for the presence of *eae* gene (Table 1).

Attempt to isolate VTEC from TSI samples

The isolation of VTEC from all the *vtx*-positive samples was not successful. The majority of samples did not yield colonies with the typical *E. coli* morphology. From the samples 13-0482, 13-0487 and 13-0489 we could select 30 colonies each, all proved negative for the presence of *vtx* genes.

Discussion

There is growing attention towards the recycling of the organic matter produced at wastewater treatment plants as TSI in agriculture. Treatment of wastewaters aims at removing pollutants of pharmaceutical, chemical and biological origin, before they are released into the surface water, to decrease the adverse impact on water ecosystems, and on animal and human health. Such treatments are usually based on physical processes such as sedimentation, activated sludge and trickling filters. The main consequence of an efficient water remediation process is the enrichment of the produced sludges with a variety of pollutants embedded in the organic matter. Such sludges usually do not undergo the removal of biological pollutants by disinfection (e.g. the use of chlorine and sometimes ozone, paracetic acid and UV irradiation) (Mezzanotte *et al.* 2007), as in the case of the wastewater effluents. For MSS the regulatory frame is mainly based on indicator organisms, such as coliforms and/or enterococci, whose load however does not always correlate with that of pathogens, such as viruses, stressed bacterial pathogens and protozoa (Fong and Lipp 2005; Okoh *et al.* 2010). The only actual pathogen considered in the current regulations is *Salmonella*, whose detection in MSS may not be exhaustive of all the potential biological hazards (Saveyn and Eder 2014).

For the above reasons, the present study aimed at investigating the presence of viral and bacterial enteric pathogens in selected TSI intended for fertilization of soils for crops and fodder production, horticulture and grazing activities. In these samples real time PCR amplification allowed the identification of the presence of sequences specific for HuNoV and adenovirus and diarrhoeagenic *E. coli* (Table 1).

pAdV and HuAdV nucleic acid detection was carried out to identify the presence of human and porcine faecal matter in the samples known to contain such matter, with the aim of correctly assigning their input source. Indeed, pAdV is ubiquitous in swine and is shed at high concentrations in faeces (Hundesa *et al.* 2009), and HuAdV are excreted in high quantities in the faeces or urine of a high percentage of human beings (Hundesa *et al.* 2006) and have already been detected in BSO (Viau and Peccia 2009). This strategy, however, was not completely successful, since we detected HuAdV in only two specimens, while none of the specimens were positive in the pAdV PCR. Four samples showed the presence of genomic traits associated with HuNoV, in two specimens in association with HuAdV. This observation may reflect a higher circulation of norovirus than adenovirus or a poor sensitivity of the real time PCR due to the presence of PCR inhibitors, revealed by failure of IAC detection in HuAdV real time PCR (two samples). However, we cannot exclude the presence of virus at concentrations below the detection limit or that the recovery efficiency of the nucleic acid following extraction may have been sub-optimal.

Interestingly, one of the four HuNoV positive samples was a MCO, which also contained human-derived sludges, confirmed by the positivity to HuAdV. The detection of HuNoV genogroups I and II poses the question of a possible risk associated with the use of MCO and BSO, since norovirus is highly infectious and even a low level of contamination could represent a threat. This result indicates that even if this particular TSI is considered at low risk of containing viruses, given the reduced amount of BSO allowed in MCO (up to 35%, under national provisions of law), its role in the transmission of such pathogens to soils and possibly to crops and vegetables should not be excluded *a priori*.

The identification of norovirus genomic sequences in BSO and MCO is not surprising, since this virus can be resistant to chemical treatments and shows a high environmental stability, over 2 months in water (Lopman *et al.* 2012). Furthermore, several studies proved that MNV, a cultivable calicivirus resembling HuNoV, can be internalized in leafy greens (Dicaprio *et al.* 2012), green onion (Hirneisen *et al.* 2012) and strawberry (DiCaprio *et al.* 2015) confirming that the use of faecally contaminated soil may pose a risk to the safety of some ready-to-eat products.

TSI were also investigated for the presence of diarrhoeagenic *E. coli*-associated virulence genes by extracting the DNA both directly from the specimens and from the cultures obtained by enriching the samples into a non-selective bacteriological medium. Such a double approach was carried out with the aim of not missing possible signals from viable but nonculturable (VBNC) microbial cells of pathogenic *E. coli* that would be outnumbered by

viable cells of other species during the enrichment. Such a possibility was considered highly possible since the VBNC is a state that bacteria can undergo when exposed to sub-lethal stresses such as those samples had likely undergone: changes in the temperature and pH, starvation, or exposure to chemicals (Liu *et al.* 2008; Oliver 2010).

Overall, we could identify the presence of virulence genes characteristic of all the pathogenic *E. coli* groups investigated in the majority of the samples (Table 1). Interestingly, the number of samples positive for the genes of EAEC and ETEC was higher when the DNA was extracted directly from the TSI samples than in the samples analysed after enrichment (nine *vs* one and four *vs* none, respectively). This result may indicate either that the strains belonging to these pathogroups are at a disadvantage in enrichment cultures, possibly due to competition with the concomitant background microflora, or that the nucleic acid amplified originated from VBNC or dead cells. Eight and seven samples were positive for the *eae* gene among the non-enriched and enriched samples, respectively. In most cases the *eae* gene was present together with *vtx* genes. Such a combination of virulence genes is generally present in VTEC strains causing severe disease in humans, such as haemorrhagic colitis and haemolytic uraemic syndrome (Caprioli *et al.* 2005), and the possible presence of these hazardous pathogens in TSI should be considered for a health risk-oriented analysis. As a matter of fact, a long term survival (more than 200 days) has been observed for VTEC of certain serogroups in manure-amended soil (van Hoek *et al.* 2013), and their internalization in plants during growth in contaminated soil has been demonstrated (Solomon *et al.* 2002; Franz *et al.* 2007).

Remarkably, in several samples we could identify the concomitant presence of *vtx* genes and virulence genes from *E. coli* pathogroups different from VTEC. Six non-enriched (five BSO and one MCO) and one enriched (BSO) sample were positive for the presence of *vtx2* and *aggR* genes, whereas two non-enriched samples were positive for *vtx2* and *ipaH* (Table 1). It has been shown that *E. coli* strains belonging to different pathogroups may acquire the bacteriophages harbouring *vtx*-genes leading to the emergence of new pathogenic variants, possessing shuffled virulence characteristics (Schmidt *et al.* 1999; Bielaszewska *et al.* 2011; Tozzoli *et al.* 2014) as happened with the VT-producing EAEC O104:H4 that caused the huge sprouts-associated outbreak in Germany and France in 2011 (Bielaszewska *et al.* 2011; Scheutz *et al.* 2011). It can be inferred that the co-occurrence of *vtx* genes, either within bacteria or conveyed by free-living phages, and pathogenic *E. coli* belonging to different pathogroups in TSI could favour the possibility of the emergence of new pathogenic VTEC types, such as VT-producing EAEC. We did not succeed in isolating the bacteria responsible for the amplification of the *vtx* genes. This could be due to the presence in the samples analysed of a strong component of non-*E. coli* flora that had outgrown the VTEC, or alternatively, to the presence in the tested samples of free *vtx*-phages that were not cultivated during the enrichment. However, the evidence of the presence of the sole virulence genes should also be considered in the risk evaluation. It cannot be excluded, in fact, that such virulence associated genetic determinants, all borne on mobile genetic elements, could be acquired by bacterial species present in the soils or the crops growing in the fields where BSO-derived TSI have been spread, resulting in the emergence of new potentially pathogenic bacterial strains.

In conclusion, the evidence presented here suggests that the routine processes leading to the stabilization of MSS, derived DIG and composts may be not fully effective in eliminating enteric viruses and diarrhoeagenic *E. coli*. This clearly indicates the need for a complete risk assessment exercise for better defining end-of-waste criteria of BSO-based TSI tailored to food safety requirements and consequently to reconsider the processes and the management options for the safe use of such valuable organic carbon sources for agriculture. Additionally, a more comprehensive approach, aimed at an unselective detection of pathogens and antimicrobial resistance genes, should be developed to fully characterize the microbial hazards possibly present in such BSO.

Acknowledgements

The present study was partially supported by the Italian Ministry of Health RF-2009-1534860 grant, project: 'ENVIFOOD: Assessment and control of selected priority contaminants intake through the characterization of rural and indoor environments'.

Conflict of Interest

The authors have no conflict of interest to declare.

References

Allard, A., Albinsson, B. and Wadell, G. (2001) Rapid typing of human adenoviruses by a general PCR combined with restriction endonuclease analysis. *J Clin Microbiol* **39**, 498–505.

Baert, L., Wobus, C.E., Van Coillie, E., Thackray, L.B., Debevere, J. and Uyttendaele, M. (2008) Detection of murine norovirus 1 by using plaque assay, transfection assay, and real-time reverse transcription-PCR before and after heat exposure. *Appl Environ Microbiol* **74**, 543–546.

Bielaszewska, M., Mellmann, A., Zhang, W., Kock, R., Fruth, A., Bauwens, A., Peters, G. and Karch, H. (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.

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

Cogger, C.G., Forge, T.A. and Neilsen, G.H. (2006) Biosolids recycling: nitrogen management and soil ecology. *Can J Soil Sci* **86**, 613–620.

Di Bartolo, I., Diez-Valcarce, M., Vasickova, P., Kralik, P., Hernandez, M., Angeloni, G., Ostanello, F., Bouwknegt, M. *et al.* (2012) Hepatitis E virus in pork production chain in Czech Republic, Italy, and Spain, 2010. *Emerg Infect Dis* **18**, 1282–1289.

Dicaprio, E., Ma, Y., Purgianto, A., Hughes, J. and Li, J. (2012) Internalization and dissemination of human norovirus and animal caliciviruses in hydroponically grown romaine lettuce. *Appl Environ Microbiol* **78**, 6143–6152.

DiCaprio, E., Culbertson, D. and Li, J. (2015) Evidence of the internalization of animal caliciviruses via the roots of growing strawberry plants and dissemination to the fruit. *Appl Environ Microbiol* **81**, 2727–2734.

Diez-Valcarce, M., Kovac, K., Cook, N., Rodríguez-Lázaro, D. and Hernández, M. (2011) Construction and analytical application of internal amplification controls (IAC) for detection of foodborne viruses by (RT) real-time PCR. *Food Anal Methods* **4**, 437–445.

EFSA (2014) EFSA BIOHAZ Panel. Scientific Opinion on the risk posed by pathogens in food of non-animal origin. Part 2 (*Salmonella*, *Yersinia*, *Shigella* and Norovirus in bulb and stem vegetables, and carrots). *EFSA J* **12**, 3937.

Fong, T.T. and Lipp, E.K. (2005) Enteric viruses of humans and animals in aquatic environments: health risks, detection, and potential water quality assessment tools. *Microbiol Mol Biol Rev* **69**, 357–371.

Franz, E., Visser, A.A., Van Diepeningen, A.D., Klerks, M.M., Termorshuizen, A.J. and van Bruggen, A.H. (2007) Quantification of contamination of lettuce by GFP-expressing *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Typhimurium. *Food Microbiol* **24**, 106–112.

Fricker, M., Messelhausser, U., Busch, U., Scherer, S. and Ehling-Schulz, M. (2007) Diagnostic real-time PCR assays for the detection of emetic *Bacillus cereus* strains in foods and recent food-borne outbreaks. *Appl Environ Microbiol* **73**, 1892–1898.

Hirneisen, K.A., Sharma, M. and Kniel, K.E. (2012) Human enteric pathogen internalization by root uptake into food crops. *Foodborne Pathogens Dis* **9**, 396–405.

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.

Hundesa, A., Maluquer de Motes, C., Bofill-Mas, S., Albinana-Gimenez, N. and Girones, R. (2006) Identification of human and animal adenoviruses and polyomaviruses for determination of sources of fecal contamination in the environment. *Appl Environ Microbiol* **72**, 7886–7893.

Hundesa, A., Maluquer de Motes, C., Albinana-Gimenez, N., Rodriguez-Manzano, J., Bofill-Mas, S., Sunen, E. and Rosina Girones, R. (2009) Development of a qPCR assay for the quantification of porcine adenoviruses as an MST tool for swine fecal contamination in the environment. *J Virol Methods* **158**, 130–135.

Jothikumar, N., Cromeans, T.L., Robertson, B.H., Meng, X.J. and Hill, V.R. (2006) A broadly reactive one-step real-time RT-PCR assay for rapid and sensitive detection of hepatitis E virus. *J Virol Methods* **131**, 65–71.

Jothikumar, N., Kang, G. and Hill, V.R. (2009) Broadly reactive TaqMan assay for real-time RT-PCR detection of rotavirus in clinical and environmental samples. JIN2@cdc.gov. *J Virol Methods* **155**, 126–131.

Kageyama, T., Kojima, S., Shinohara, M., Uchida, K., Fukushi, S., Hoshino, F.B., Takeda, N. and Katayama, K. (2003) Broadly reactive and highly sensitive assay for Norwalk-like viruses based on real-time quantitative reverse transcription-PCR. *J Clin Microbiol* **41**, 1548–1557.

Lgs, D. (2010) Decreto Legislativo 29 aprile 2010, n. 75 "Riordino e revisione della disciplina in materia di fertilizzanti, a norma dell'articolo 13 della legge 7 luglio 2009, n. 88" , Gazzetta Ufficiale n. 121 del 26 maggio 2010.

Lgs, D. (2013) Decreto 10 luglio 2013 Ministero delle politiche agricole alimentari e forestali. Aggiornamento degli allegati del decreto legislativo 29 aprile 2010, n. 75, concernente il riordino e la revisione della disciplina in materia di fertilizzanti. Gazzetta Ufficiale n.218 del 17 settembre 2013).

Liu, Y., Gilchrist, A., Zhang, J. and Li, X.F. (2008) Detection of viable but nonculturable *Escherichia coli* O157:H7 bacteria in drinking water and river water. *Appl Environ Microbiol* **74**, 1502–1507.

Liu, J., Gratz, J., Amour, C., Kibiki, G., Becker, S., Janaki, L., Verweij, J.J., Taniuchi, M. *et al.* (2013) A laboratory-developed TaqMan Array Card for simultaneous detection of 19 enteropathogens. *J Clin Microbiol* **51**, 472–480.

Lopman, B., Gastanaduy, P., Park, G.W., Hall, A.J., Parashar, U.D. and Vinje, J. (2012) Environmental transmission of norovirus gastroenteritis. *Curr Opin Virol* **2**, 96–102.

Maluquer de Motes, C., Clemente-Casares, P., Hundesa, A., Martin, M. and Girones, R. (2004) Detection of bovine and porcine adenoviruses for tracing the source of fecal contamination. *Appl Environ Microbiol* **70**, 1448–1454.

Mezzanotte, V., Antonelli, M., Citterio, S. and Nurizzo, C. (2007) Wastewater disinfection alternatives: chlorine,

ozone, peracetic acid, and UV light. *Water Environ Res* **79**, 2373–2379.

Nielsen, E.M. and Andersen, M.T. (2003) Detection and characterization of verocytotoxin-producing *Escherichia coli* by automated 5' nuclease PCR assay. *J Clin Microbiol* **41**, 2884–2893.

NRC (2002) *Biosolids Applied to Land Advancing Standards and Practices Committee on Toxicants and Pathogens in Board on Environmental Studies and Toxicology, Division on Earth and Life Studies*. Washington, DC: The National Academies Press.

Okoh, A.I., Sibanda, T. and Gusha, S.S. (2010) Inadequately treated wastewater as a source of human enteric viruses in the environment. *Int J Environ Res Public Health* **7**, 2620–2637.

Oliver, J.D. (2010) Recent findings on the viable but nonculturable state in pathogenic bacteria. *FEMS Microbiol Rev* **34**, 415–425.

Perelle, S., Dilasser, F., Grout, J. and 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.

Rhodes, E.R., Boczek, L.A., Ware, M.W., McKay, M., Hoelle, J.M., Schoen, M. and Villegas, E.N. (2015) Determining pathogen and indicator levels in class B municipal organic residuals used for land application. *J Environ Quality* **44**, 265–274.

Saveyn, H. and Eder, P. (2014) *End-of-Waste Criteria for Biodegradable Waste Subjected to Biological Treatment (Compost & Digestate): Technical Proposals*. Sevilla, Spain: IPTS/EU Commission.

Scheutz, F., Nielsen, E.M., Frimodt-Moller, J., Boisen, N., Morabito, S., Tozzoli, R., Nataro, J.P. and Caprioli, A. (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 Surveillance Eur Commun Dis Bull* **16**, pii: 19889.

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.

da Silva, A.K., Le Saux, J.C., Parnaudeau, S., Pommepuy, M., Elimelech, M. and Le Guyader, F.S. (2007) Evaluation of removal of noroviruses during wastewater treatment, using real-time reverse transcription-PCR: different behaviors of genogroups I and II. *Appl Environ Microbiol* **73**, 7891–7897.

Solomon, E.B., Yaron, S. and Matthews, K.R. (2002) Transmission of *Escherichia coli* O157:H7 from contaminated manure and irrigation water to lettuce plant tissue and its subsequent internalization. *Appl Environ Microbiol* **68**, 397–400.

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.

Usman, K., Khan, S., Ghulam, S., Khan, M.U., Khan, N., Khan, M.A. and Khalil, S.K. (2012) Sewage sludge: an important biological resource for sustainable agriculture and its environmental implications. *Am J Plant Sci* **3**, 1708–1721.

Viau, E. and Peccia, J. (2009) Survey of wastewater indicators and human pathogen genomes in biosolids produced by class a and class B stabilization treatments. *Appl Environ Microbiol* **75**, 164–174.

Viau, E., Bibby, K., Paez-Rubio, T. and Peccia, J. (2011) Toward a consensus view on the infectious risks associated with land application of sewage sludge. *Environ Sci Technol* **45**, 5459–5469.

Wang, S.M., Ma, J.C., Hao, Z.Y., Zhang, Z.Y., Mason, C., Sethabutr, O., von Seidlein, L., Wang, X.Y. *et al.* (2010) Surveillance of shigellosis by real-time PCR suggests underestimation of shigellosis prevalence by culture-based methods in a population of rural China. *J Infect* **61**, 471–475.

Wiechman, B., Dienema, C., Kabbe, C., Brandt, S., Vogel, I. and Roskosch, A. (2015) *Sewage Sludge Management in Germany*. Dessau-Roβlau, Germany: Umweltbundesamt (UBA).

Chapter 5

Comparative analysis of metagenomes of Italian top soil improvers

5. Comparative analysis of metagenomes of Italian top soil improvers

In the previous Chapter, we have shown that the use of Biosolids (BSO) to amend agricultural land may present the risk of spreading known or emerging hazards of anthropogenic or animal origin in the environment devoted to food production. We described the presence of virulence genes associated with pathogenic *E. coli*, including STEC, and enteric viruses in BSO samples. These findings suggested the need to perform a holistic and unbiased analysis of such BSO. For this purpose, we used a metagenomic approach to determine a deep microbiological profiling of BSO specimens derived from both sludges and compost. The sludge used to prepare the BSO were from wastewater treatment plants collecting both urban settlements and livestock wastes, while the compost based-BSO were mainly from green wastes.

The taxonomic classification of the microbiota present in the BSO samples allowed associating the taxonomic composition of the specimens'microbiota with the origin of the samples. In particular, we could clearly separate the specimens collected from sludges-based BSO from those resulting from compost. Our results indicated that the microbial content of different BSO can be a good indicator of the samples' origin and that the analytical approach used can be an efficacious tool to trace their origin.

Additionally, the analysis showed the presence of many genomics traits associated with different diarrheagenic *E. coli* pathogroups as well as the presence of determinants conferring resistance to antimicrobials (AM) and other compounds. In particular, the metagenomic screening highlighted the presence of genes associated with the resistance to the heavy metals, and to the main classes of antibiotics, as Fluoroquinolones, Beta-lactams, Streptothricin, Fosfomycin, Vancomycin and Methicillin. These findings suggest that amending soils intended for food production with BSO may introduce the possibility to transfer antimicrobial resistance (AMR) genes to the soil and the crop bacterial communities, contributing to the overall spread of antimicrobial resistance.

Finally, the holistic approach used allowed us to identify a high number of combinations of virulence genes of pathogenic *E. coli* in the BSO, these results also confirm what it has been already shown with other approaches in Chapter 4, strengthening the need for a thorough risk assessment of the use of such soil improvers. This is particularly true when STEC are considered. As a matter of fact, particularly when the BSO are obtained from sludges from urban and zootechnical wastes the overlapping of STEC from human and animal sources may favour the exchange of MGE, including the Stx-phages, and facilitate the emergence of strains with shuffled virulence determinants.

Environmental Research 155 (2017) 108–115

Contents lists available at ScienceDirect

Environmental Research

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

Comparative analysis of metagenomes of Italian top soil improvers

Federica Gigliucci[a,b,*], Gianfranco Brambilla[a], Rosangela Tozzoli[a], Valeria Michelacci[a], Stefano Morabito[a]

[a] Department of Veterinary Public Health and Food Safety, Istituto Superiore di Sanità, Viale Regina Elena, 299 00161 Rome, Italy
[b] Department of Sciences, University Roma,Tre, Viale Marconi, 446, 00146 Rome, Italy

ARTICLE INFO

Keywords:
Top soil improvers
Metagenomics
Microbiome
Pathogenic *E. coli*
Antimicrobial resistance

ABSTRACT

Biosolids originating from Municipal Waste Water Treatment Plants are proposed as top soil improvers (TSI) for their beneficial input of organic carbon on agriculture lands. Their use to amend soil is controversial, as it may lead to the presence of emerging hazards of anthropogenic or animal origin in the environment devoted to food production. In this study, we used a shotgun metagenomics sequencing as a tool to perform a characterization of the hazards related with the TSIs. The samples showed the presence of many virulence genes associated to different diarrheagenic *E. coli* pathotypes as well as of different antimicrobial resistance-associated genes. The genes conferring resistance to Fluoroquinolones was the most relevant class of antimicrobial resistance genes observed in all the samples tested. To a lesser extent traits associated with the resistance to Methicillin in *Staphylococci* and genes conferring resistance to Streptothricin, Fosfomycin and Vancomycin were also identified. The most represented metal resistance genes were cobalt-zinc-cadmium related, accounting for 15–50% of the sequence reads in the different metagenomes out of the total number of those mapping on the class of resistance to compounds determinants. Moreover the taxonomic analysis performed by comparing compost-based samples and biosolids derived from municipal sewage-sludges treatments divided the samples into separate populations, based on the microbiota composition. The results confirm that the metagenomics is efficient to detect genomic traits associated with pathogens and antimicrobial resistance in complex matrices and this approach can be efficiently used for the traceability of TSI samples using the microorganisms' profiles as indicators of their origin.

1. Introduction

The recycling of bio-waste represents a resource for energy, water and nutrients in agriculture (ISWA, 2013). The bio-waste category includes sewage sludges derived from biological or chemical treatment of industrial or municipal wastewater and from the agri-food sector; manure from livestock and compost from both green wastes and organic fraction of the household wastes (Saveyn and Eder, 2014).

The solid organic matter from sewage treatment plants is often defined as Biosolid (BSO). BSO can be stabilized through an aerobic fermentation process by which mesophilic and thermophilic micro-organisms decompose organic matter into simpler nutrients. The process consumes oxygen and produces heat, with temperature rising up to 60 °C (Liang et al., 2003; Trautmann and Olynciw, 2010).

In the last years an increase in the use of BSO from Municipal Waste Water Treatment Plants (WWTP) in agriculture as top soil improvers (TSI) (1,000,000 t/year in Italy, mean 5 t/hectare/year) was observed (ISPRA, 2012). Such a trend relates to the presence of useful compounds of potential environmental value, such as organic carbon, nitrogen, phosphorous and potassium and to lesser extent, calcium, sulphur and magnesium (Usman et al., 2012). This approach allows facing intensive cropping while reducing the need to make use of the more expensive mineral fertilizers, making the use of BSO sustainable and economical due to nutrient cycling and disposal of sewage sludge (Usman et al., 2012).

In contrast to the proposed benefits, the use of BSO as TSIs has drawbacks. It has been shown that BSO and the derived TSIs can be 10–100 fold more contaminated with persistent organic pollutants (POPs) than the animal manure (Brambilla et al., 2016). The use of such products to improve the fields fertility may cause the soils to be persistently contaminated in the range of the hundreds or thousands of ng/g dry matter for certain POPs (Gottschall et al., 2012; Zennegg

Abbreviations: BSO, Biosolid; WWTP, Municipal Waste Water Treatment Plants; TSI, Top Soil Improver; POP, Persistent Organic Pollutants; STEC, Shiga Toxin producing *E. coli*; CO, Compost; MCO, Mixed Compost; AMR, Antimicrobial Resistance; COG, Cluster of Orthologous Groups; Kegg, Kyoto Encyclopaedia of Genes and Genomes; PCoA, Principal Coordinate Analysis; EAEC, Enteroaggregative *E. coli*; EPEC, Enteropathogenic *E. coli*

* Corresponding author at: Dipartimento di Sanità Pubblica Veterinaria e Sicurezza Alimentare, Istituto Superiore di Sanità, Viale Regina Elena, 299 00161 Rome, Italy.
E-mail address: Federica.gigliucci@libero.it (F. Gigliucci).

http://dx.doi.org/10.1016/j.envres.2017.02.004
Received 15 December 2016; Received in revised form 2 February 2017; Accepted 7 February 2017
Available online 10 March 2017
0013-9351/

et al., 2013; Suominen et al., 2014). The top soil intake by grazing animals represents the main determinant for the chemical carry-over to food and exposure through food consumption in humans (Brambilla et al., 2011). Beside the chemical contamination, the use of TSIs of anthropogenic and farm origin may cause the spreading of zoonotic and human enteric pathogens in pastures (Yergeau et al., 2016). Recently, virulence genes associated to important human pathogens, such as Shiga Toxin producing *E. coli* (STEC) and enteric viruses, were detected by Real Time PCR in TSIs (Tozzoli et al., 2016). A recent study shown that biosolids contain numerous virulence-associated genes related with many pathogens as assessed by metagenomics shotgun sequencing (Yergeau et al., 2016). It is noteworthy that a long term survival (more than 200 days) has been observed for STEC O157 in manure-amended soil (van Hoek et al., 2013). Noro- and Rota-viruses are resistant to wastewater treatment and can persist in the raw sewage and run-off waters from WWTPs as well as in surface waters used in agriculture for several months (López-Gálvez et al., 2016; Zhou et al., 2016). Similarly, parasites as *Cryptosporidium parvum* and *Giardia lamblia* have been shown to persist in soils for long times (Helmi et al., 2008). Besides the possibility of transmission of microbial pathogens, transfer of antimicrobial resistance genes from manure to soil bacteria has been described (Heuer and Smalla, 2007; Binh et al., 2008). Sengeløv et al. (2003) reported that resistance to Tetracycline, Macrolides and Streptomycin was measured for a period of 8 months in soil bacteria obtained from farmland treated with pig manure slurry (Sengeløv et al., 2003). A recent study on the antibiotic resistance in sewage treatment plants has shown that antimicrobial resistance genes can be enriched during sludge treatment process (Bengtsson-Palme et al., 2016).

We used a shotgun metagenomics sequencing approach, to perform the characterization of biological hazards related to the use of TSIs in agriculture in Italy. We selected eight samples including BSO, compost (CO) and mixed compost (MCO) and determined the microbiological profile diversity and to assess the presence of genomics traits related with pathogenic *E. coli* and antimicrobial resistance (AMR) determinants.

2. Materials and methods

2.1. Sampling and samples origin

The specimens analysed in this study were selected among a collection of 24 Top soil improvers samples collected in 2013 from different Italian regions and used in a previous study (Tozzoli et al., 2016). In particular, eight of these have been used for the metagenomics analysis: four municipal sewage sludges (BSO1, BSO2, BSO3, BSO4), obtained from a tertiary wastewater treatment plants with merged inputs from urban and intensive farmed pigs settlements; three samples of compost, (CO1, CO2 and CO3) derived from green wastes only; and one mixed compost sample (MCO) containing contributions from household wastes, green wastes and urban sewage.

Sampling was performed in accordance with UNI 10802/2004 guideline and the provision of law for the analysis of heavy metals. Biosolids from Waste Water Treatment Plants were stored in glass jars, in the dark, at − 30 °C, until the analyses were performed.

2.2. Nucleic acid extraction and DNA sequencing

DNA was extracted from 0.25 g of each untreated sample using the Power Soil DNA isolation kit (MO BIO Laboratories inc., Carlsbad, CA, USA) following the manufacturer's instructions.

Sequencing libraries were prepared from 100 ng of the DNA extracted from each sample, using the NEBNext Fast DNA Fragmentation & Lirbary Prep kit (New England BioLabs, New England, USA). In detail, the DNA was enzymatically fragmented to obtain fragments of about 400 bp, through an incubation at 25 °C for 15 min, followed by 10 min at 70 °C. The fragmented DNA was subjected to link with adaptors and size selection of 450 bp fragments by electrophoresis on E-Gel SizeSelect 2% (Invitrogen, Carlsbad, USA) followed by PCR amplification as indicated in the NEBNext Fast DNA Fragmentation & Library Prep kit manual (New England BioLabs, USA). The libraries were amplified individually through emulsion PCR with an Ion OneTouch 2 and sequenced with an Ion Torrent Personal Genome Machine (Life Technologies, 118 Carlsbad, USA), using the 400 bp sequencing protocol. The eight samples were sequenced individually in eight different runs using a 316 V2 chip per run.

2.3. Bioinformatics analysis

The sequences were analysed using the open source webserver for processing metagenomics sequences data MG-RAST to characterize the relative microbiota composition (http://metagenomics.anl.org) (Meyer et al., 2008). In detail, the sequence data were filtered for the sequences of human origin and analysed for the microbial content and then the resulting reads were compared against M5NR (the M5 non-redundant) protein database, using a maximum e-value of 1e-5, a minimum identity of 60%, and a minimum alignment length of 15 aa for protein and 15 bp for RNA databases (Meyer et al., 2008).

The bioinformatics analyses of the metagenomes were also performed using the tools available on the ARIES public webserver (https://w3.iss.it/site/aries/).

In detail, the raw reads were subjected to a quality check and consequent trimming to remove the adaptors and to accept 20 as the lowest Phred value. The identification of the presence of *E. coli* virulence genes was performed through the pipeline Virulotyper, which employs Bowtie2 algorithm (http://bowtiebio.sourceforge.net/bowtie2/) (Langmead and Salzberg, 2012) to map the sequencing reads against the *E. coli* Virulence genes database (Joensen et al., 2014). Virulence genes showing coverage above 1X were considered present in the sample.

Comparative analysis of all samples was performed by using COMMET (COmpare Multiple METagenomes) (Maillet et al., 2014). This tool allows to observe the homology between the different samples on the basis of all-against-all comparisons of the non-assembled reads (http://colibread.inria.fr/commet/) (Maillet et al., 2014).

Functional metagenomics analysis was performed using the DIAMOND tool (Buchfink et al., 2015). Sequence data in a FASTA format were used to search the COG (Cluster of Orthologous Groups) reference database (Tatusov et al., 1997).

The results of the DIAMOND alignments were analysed and visualized using the MEGAN (MEta Genome ANalyzer) software version 5 (http://ab.inf.uni-tuebingen.de/software/megan5/) (Huson et al., 2007). The SAM files produced with the DIAMOND tool were imported into MEGAN 5 together with the FASTA files containing the sequence data. The KEGG (Kyoto Encyclopaedia of Genes and Genomes) (Kanehisa and Goto, 2000), COG and SEED subsystems (Overbeek et al., 2005) content was determined using the MEGAN internal Reference sequence maps. MEGAN 5 software was also used to perform the rarefaction analysis (Gotelli and Colwell, 2001) and to calculate the distances between the different samples through the principal coordinate analysis (PCoA) (Smith et al., 2007). For the latter analysis a stress of 0.41 and the Bray Curtis ecological index were used.

3. Results

3.1. Metagenomes quality check

On average, 2,861,113 reads per sample were retained after the quality check and have been used for the subsequent analyses.

The metagenomics datasets are available at the MG-RAST website, under the identification numbers: 4639314.3 (CO1); 4639313.3 (CO2); 4639304.3 (CO3); 4631825.3 (MCO); 4633025.3 (BSO1); 4632987.3

(BSO2); 4631959.3 (BSO3); 4631832.3 (BSO4).

3.2. Taxonomic profiling and microbial diversity

Rarefaction analysis allows assessing the species richness in a metagenome allowing for meaningful standardization and comparison of datasets, based on the construction of the rarefaction curves (Gotelli *et al.*, 2001). Such an analysis was carried out with the MEGAN 5 software and showed that the profiling of the metagenomes of the compost samples produced a lower number of leaves in the taxonomy tree when compared to the sludges (see Supplementary materials, Fig. S1), indicating a lower abundance in the taxonomic units in the former samples.

The metagenomes were then analysed at the MG-RAST webserver for the taxonomic analysis. The sequences were filtered for removing the DNA sequences of human origin and analysed for the microbial content. The MG-RAST result showed an evident predominance of Bacteria vs Archaea, the former representing 92–98% of the total microorganisms' content and the latter only accounting for 0.5–5.9%. Approximately, 1.5% of the organisms were from the Eukaryota domain in all the samples, while the remaining comprised DNA viruses and unclassified organisms.

Euryarchaeota and Crenarchaeota constituted most of archaeal taxa in each analysed TSI specimen, without a specific correlation with their source. As for the Eukaryota domain, a possible relation was observed between the presence of members of the Ascomycota phylum and the origin of the TSI samples. As a matter of fact, the compost-based samples CO1, CO2 and CO3 exhibited a high content of Ascomycota, forming the most represented Eukaryotic taxon (80–90%), while their relative abundance was reduced in the MCO (39%) and even lower in all sludges (around 16%).

Actinobacteria (2.6–40%), Proteobacteria (12.8–40%), Bacteroidetes (6.3–65.5%) and Firmicutes (3.6–45.2%) were the most abundant phyla of Bacteria, whereas the phylum Chloroflexi showed a good abundance (14.7%) only in one sample (Table 1).

In addition to Ascomycota, also some orders and classes of the bacterial phyla Proteobacteria, Firmicutes, Chloroflexi and Actinobacteria showed a distribution in the metagenomes markedly associated with the samples' origin. In particular, the order of Burkholderiales and the classes Bacilli and Thermomicrobia were significantly more represented among the compost samples ($P < 0.0001$), while the order of Rhodocyclales and the classes of Clostridia and Anaerolineae were more represented in the sludge samples ($P < 0.0001$) (Fig. 1).

Accordingly, the comparative analyses performed with the PCoA analysis clearly divided the samples into well separate populations (compost vs sludges), based on the composition of the microbiota (Fig. 2B). Similarly, the topology of the dendrogram generated with COMMET showed that the samples grouped under two main branches identifying the sludges population with respect to the compost samples analysed (Fig. 2A).

3.3. Virulence genes associated to pathogenic E. coli

As the TSI samples analysed in this study were already described to contain genomic traits associated with pathogenic *E. coli* (Tozzoli et al., 2016), we assayed the capability of the metagenomics approach to identify the presence of such genes in the metagenomes. The mapping of the trimmed raw reads against the *E. coli* Virulence genes database (Joensen et al., 2014) showed the presence of many virulence genes associated to different diarrheagenic *E. coli* pathotypes (Table 2). This analysis showed that the metagenomics approach can be as effective as other established technologies, such as the real time PCR, in detecting *E. coli* virulence genes associated with the different *E. coli* pathogroups. As a matter of fact, our analysis allowed us, in most cases, to identify genes associated with the *E. coli* pathogroups detected in the previous study (Tozzoli et al., 2016) (Table 2). In samples MCO, BSO2, BSO3 and BSO4, the metagenomics approach failed to highlight the presence of Enteroaggregative *E. coli* (EAEC) gene *aggR*, whose presence was previously detected by real time PCR (Table 2). In the latter two samples, however, other genes related with EAEC were identified (*pic* and *pet*; Table 2). On the other hand, in sample MCO the metagenomics approach allowed us to identify the presence of the gene *ipaH*, which was also searched with negative results by real time PCR (Tozzoli et al., 2016). Moreover, through the metagenomics approach, we could observe the presence of genes associated with Enteropathogenic *E. coli* (EPEC) and Shiga-toxin producing *E. coli* (STEC) (*toxB, katP* and *tir*) in samples CO2 and CO3, which were negative for the presence of the assayed EPEC and STEC-associated genes in the previous real time PCR experiments (Tozzoli et al., 2016) (Table 2).

3.4. Genomic traits related to resistance to compounds

The functional analysis applied to the metagenomes, carried out with both the MG-RAST and the MEGAN5 software, revealed the presence, in these samples, of several genetic traits encoding resistance to compounds including antimicrobials (AM) (Fig. 3). The most represented resistance genes were those encoding the multi-drugs efflux pumps and the resistance to cobalt-zinc-cadmium, the latter accounting for 15–50% of the sequence reads in the different metagenomes out of the total number of those mapping on the class of resistance to compounds determinants. Among the antibiotics, the most relevant class of AM resistance genes (AMR) observed in all the samples tested was that conferring resistance to Fluoroquinolones, with amounts of mapped reads ranging from 16% to 26% in different metagenomes (Fig. 3). The sequences mapping on the beta-lactamases also had a good relative abundance (Fig. 3). Interestingly, the *blaR1* gene, encoding the β-lactam sensor/signal transducer in methicillin-restistant *Staphilococcus aureus* (Hao et al., 2012) was abundant in samples CO2 and BSO2, (not shown). Finally, the products of the genes conferring resistance to Streptothricin, Fosfomycin and Vancomycin were also identified, although to a lesser extent (0–1.7%) (Fig. 3). Finally, traits associated with the resistance to Methicillin in Staphylococci accounted for 2.8–7.2% of the reads (Fig. 3).

4. Discussion and conclusion

The use of TSIs represents a great benefit for the agricultural sector, as they provide valuable carbon sources and nutrients. Moreover, the recycling of the organic matter in form of sludges produced at waste-

Table 1
Relative abundance of the main bacterial phyla in the analysed TSI samples. The values indicate the percentage of reads mapping against specific traits of the considered phyla present on the M5NR database. The values depend on the total of the analysed reads.

	CO1	CO2	CO3	MCO	BSO1	BSO2	BSO3	BSO4
Actinobacteria	40%	25.70%	11.30%	13.60%	4.20%	5.50%	6.70%	2.60%
Proteobacteria	26%	12.80%	37.80%	30.70%	40%	28.40%	39.60%	17.40%
Bacteroidetes	17%	11.60%	39.50%	37.10%	25%	6.30%	27.40%	65.50%
Firmicutes	6.70%	45.20%	3.60%	4.70%	9.30%	23%	8.80%	3.80%
Chloroflexi	2.40%	0.50%	1.10%	2.60%	3.90%	14.70%	1.90%	1.80%

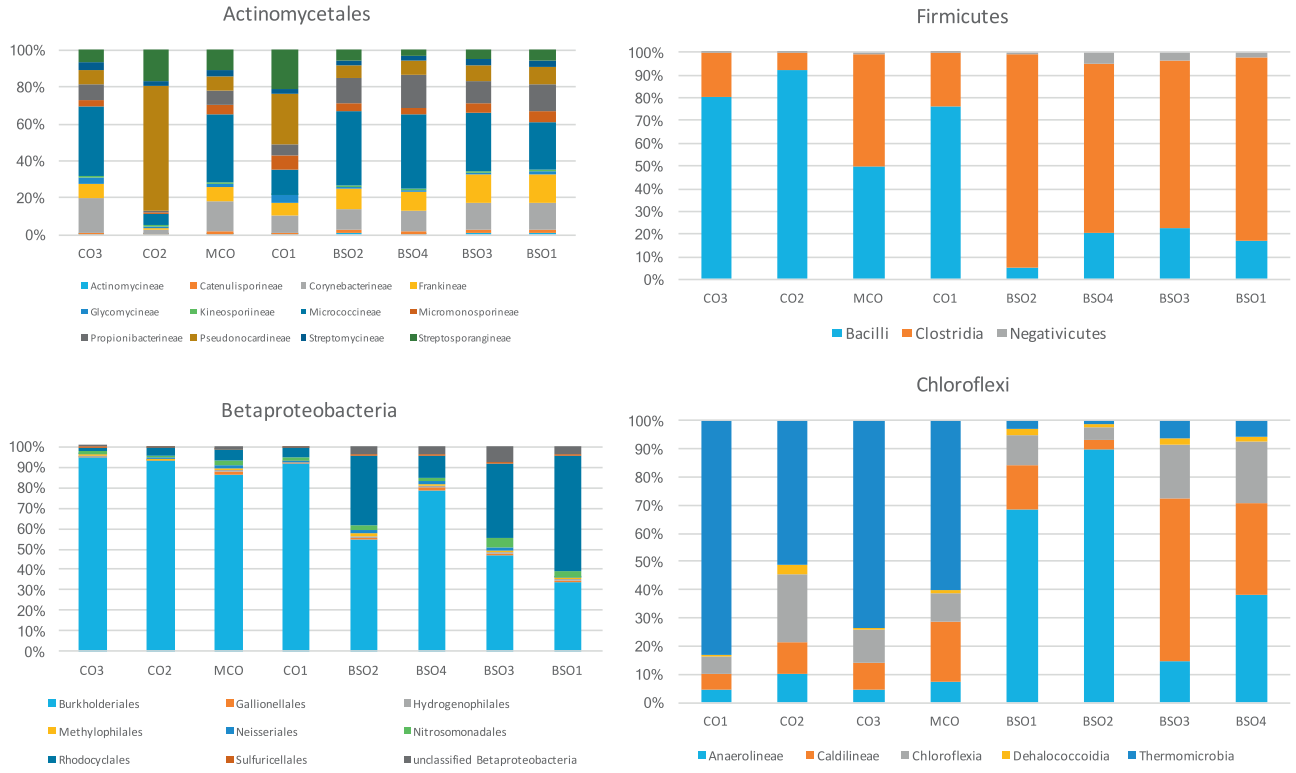


Fig. 1. Distribution of the OTUs in the metagenomes significantly associated with the samples' origin. The scale on the y axis refers to the percent of the reads mapping on the specific OTU. For all the OTUs shown the difference observed was statistically significant with a P value < 0.0001.

water treatment plants is seen as the key to improve soil fertility while decreasing the size of the wastes originating from urban settlements as well as from industrial and zootechnical settings. However, treatment of sludges, including disinfection, is generally not performed (Mezzanotte et al., 2007), thus exposing the crops grown in fields where sludge-based TSIs are used to potential risks of chemical, pharmacological, and microbial contamination of the food chain (Saveyn and Eder, 2014). In Europe a harmonized legislative framework on the subject is not in place. Therefore, the regulations vary considerably from country to country. The European Union is inves-

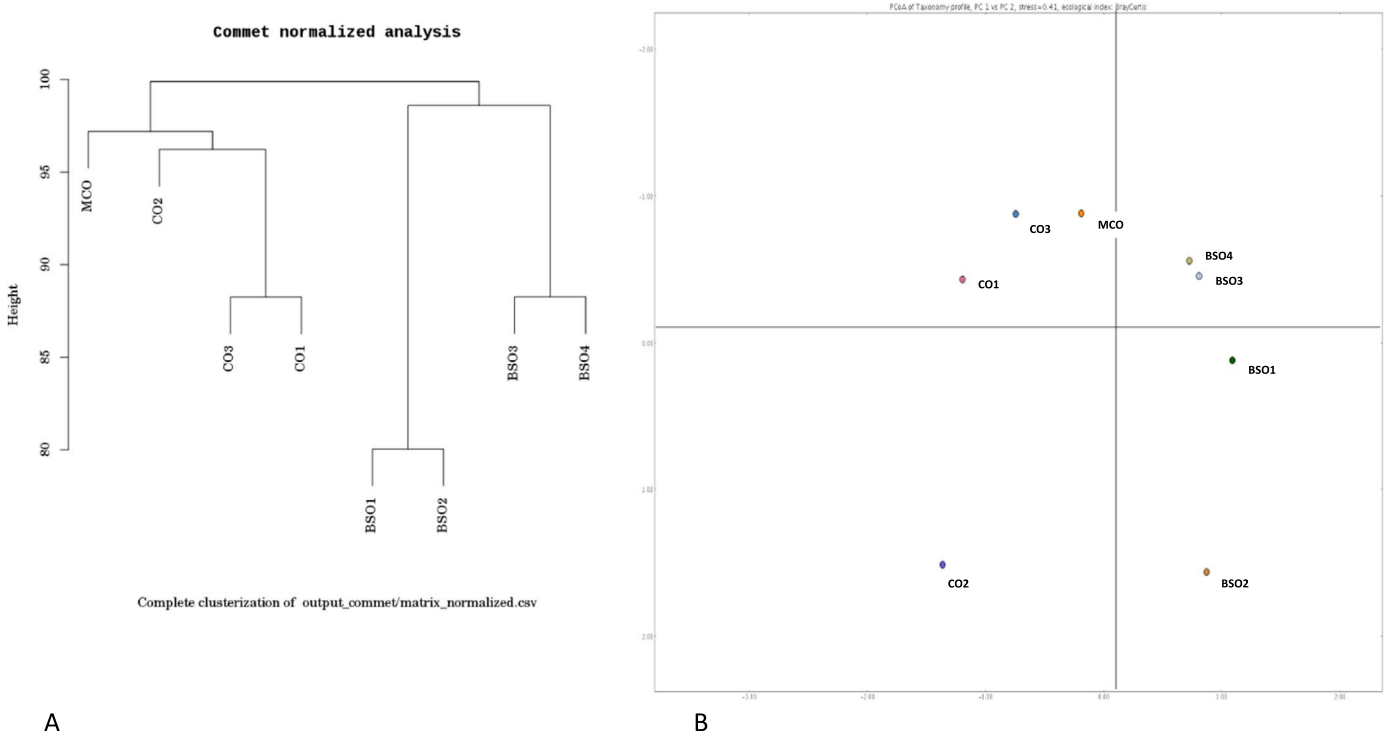


Fig. 2. Panel A: Dendrogram showing the distribution of the samples according to their origin obtained with the COMMET tool. Panel B: Principal Coordinate Analysis (PCoA), obtained by using MEGAN5 software. The localization in the graph of each sample is represented by a coloured circle and the distances between the circles reflect the original distance between the samples, calculated using the BrayCurtis ecological index. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Comparison of diarrheagenic *E. coli* virulence genes emerged by applying metagenomics approach (this study) and by real time PCR (Tozzoli et al., 2016). In bracket are reported the samples codes used in Tozzoli et al., 2016 for cross reference.

CO1 (13–0487)		CO2 (13–0485)		CO3 (13–0483)		MCO (13–0486)		BSO1 (13–0493)		BSO2 (13–0489)		BSO3 (13–0492)		BSO4 (13–0490)	
This study	**Tozzoli et al.**	**This study**	**Tozzoli et al.**	**This study**	**Tozzoli et al.**	**This study**	**Tozzoli et al.**	**This study**	**Tozzoli et al.**	**This study**	**Tozzoli et al.**	**This study**	**Tozzoli et al.**	**This study**	**Tozzoli et al.**
		espI *iha* *prfB* *tir* *toxB*		*katP* *nleB* *pet* *prfB* *tir*		*espC* *fedF* *ipaH* *prfB* *tir*	*aggR*	*katP* *prfB* *tir* *vat*	*lt*	*iroN* *prfB* *tir* *toxB*	*aggR* *stx2*	*katP* *nleB* *pet* *prfB* *tir*	*aggR* *eae* *ipaH* *stx2*	*astA; bfpA* *cba; ccl* *cdtB; cfaC* *cif; cma* *cnf1; eae* *eatA; efa1* *espA; espB* *espC; espF* *espI; espJ* *espP; etpD* *fim41A;* *ehxA* *hlyE; ihaA* *ireA; iroN* *ipaH; iss* *K88; katP* *nleB; nleC* *pet; pic* *picU; prfB* *senB;sepA* *stx1Ac;* *stx1Bc* *stx1Ad;* *stx1Bd* *stx1Aa* *stx2Aa;* *stx2Ad* *stx2Ae;* *stx2Af* *stx2Ag* *subA; saa* *tccP; tir* *toxB; tsh* *virF*	*aggR* *eae* *stx2*

tigating the definition of End of Waste criteria to aid the decision on whether certain sludges can be used as fertilizers and cease to be classified as waste (ISWA, 2013). In some European countries the biowastes used as TSIs are not allowed to contain sludge, while in others it can be used only after composting and it must fall within the scope of the End of Waste criteria for biodegradable waste, according to Article 6 of the Waste Framework Directive (2008/98/EC, Dir.).

In Italy, criteria related to the sewage sludge use in agriculture have been laid down in the Regulation 99 (99/1992, D.lgs), representing the adoption of the Sewage Sludge Directive 86/278/EEC (86/278/EEC, Dir.). The Directive defines some chemical, physical and biological parameters that TSIs must comply with, such as the absence of Salmonella and limits for *E. coli* (99/1992, D.lgs). Additionally, it states that the use of untreated sludge (e.g. by heat treatment and/or long-term storage) on agricultural land is not allowed (99/1992, D.lgs) and that these should be labelled differently from those from green and/or household wastes (10/07/2013, Dm).

In a previous study, we have described the presence of virulence genes of pathogenic *E. coli* and enteric viruses in TSIs in Italy (Tozzoli et al., 2016) and came to the conclusion that a comprehensive risk assessment is needed to come to a sound scientific basis for developing regulations.

In the present study we aimed at using metagenomics as a tool to perform a characterization of the hazards related with the TSIs, not biased towards the search of one specific threat. We have used the shotgun metagenomics to identify the presence of pathogenic *E. coli* associated genes as well as of determinants associated with resistance to compounds. Moreover, we performed the microbiological profiling of the TSIs in order to determine if the composition of the microbiota may be used to trace the origin of the matter the TSIs are made with.

The search for the virulence genes in the TSI samples through metagenomics proved effective in detecting evidence of the presence of pathogenic *E. coli*. The samples analysed in this study had been previously shown to contain genes associated with different *E. coli* pathogroups by means of a real time PCR screening (Tozzoli et al., 2016). The comparison of the metagenomics results with those obtained in the previous study indicated the former approach as a promising strategy for the assessment of the presence of genomic traits associated with any pathogen in a complex matrix, such as the TSIs, with the only limitation of the availability of comprehensive databases of markers for the pathogens of interest. As a matter of fact, our results showed a good correlation between the virulence genes content of the TSIs identified in this study and the previously published analysis (Tozzoli et al., 2016) (Table 2). In some cases, as for the samples MCO, BSO2, the metagenomics approach failed to identify genes associated with Enteroaggregative *E. coli* (EAEC). However, in samples BSO3 and BSO4, the evidence of EAEC was provided by the presence of other genes associated with this *E. coli* pathogroup, but different from those used in the real time PCR experiments described previously. This discrepancy is not unexpected as the real time PCR is based on the amplification of specific DNA sequence, which can be present in the sample in a very low abundance and below the sensitivity limit of a non-target approach as the metagenomics. On the other hand, this approach provided evidence of genes characteristic of other pathogenic *E. coli*, which were not detected with the PCR-based approach (Table 2). While the real time PCR approach may be more sensitive,

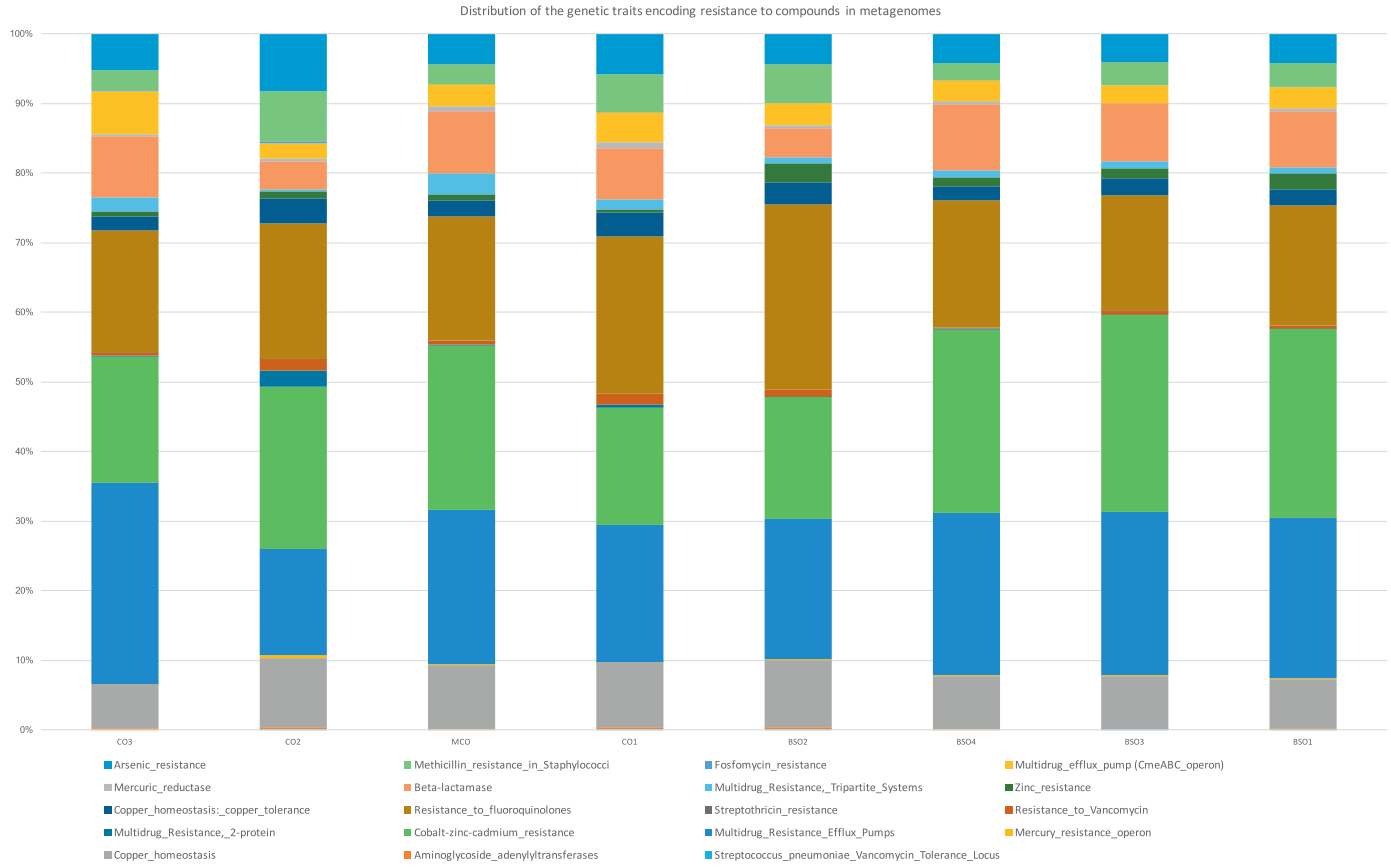


Fig. 3. Distribution of the reads mapping on specific determinants conferring resistance to compounds. The values on the Y axis correspond to the percentage of the reads mapping on specific determinants out of the total number of reads mapping on this specific class of determinants.

due to the exponential amplification of the targets, the metagenomics approach may complement the lower technical sensitivity with the simultaneous search for more determinants associated with the same pathogroup, as in the case of samples BSO3 and BSO4. It has to be stressed that the focus on the identification of pathogenic *E. coli* was chosen as a good database is available containing all the known sequences of the virulence genes and their alleles (Joensen et al., 2014). While the existence of accurately curated databases with the genetic markers for other pathogenic agents may not be available yet, this is a field of development where many research groups are currently operating. Moreover, also in the lack of specific databases, there is already the possibility to screen metagenomes for several pathogens using the functional metagenomics approach based on the detection of the virulence-associated proteins present in the large non-redundant proteins (Pruitt et al., 2005) or the seed (Overbeek et al., 2005) databases. Such an approach was already proved functional in detecting virulence genes of multiple pathogens in samples from wastewater treatment facilities in Canada (Yergeau et al., 2016).

The metagenomics approach allowed us to investigate the presence, in TSIs, of determinants conferring resistance to compounds including antimicrobials (AM). The most abundant factors identified were the multidrug-resistance efflux pumps (Fig. 3), which can confer resistance to clinically relevant antibiotics. Those pumps transport several substances produced by the host, improving resistance to host defence molecules, and can transport antibiotics of different chemical classes, conferring decreased susceptibility to antimicrobial (Piddock, 2006). The analysis also showed that genes conferring resistance to specific classes of antibiotics, such as the Fluoroquinolones and the Beta-lactams, as well as Streptothricin, Fosfomycin, Vancomycin and Methicillin, were also represented, suggesting that, following the use of these TSIs, they could be transferred to microbial soil communities, contributing to the overall spread of antibiotic resistance. Our results are in line with those reported in a recent study, describing the antimicrobial resistance genes in sewage from wastewater treatment plants in Sweden (Bengtsson-Palme et al., 2016). One of the most abundant classes of genes governing the resistance to compounds identified in the metagenomes in this study was that related with the resistance to heavy metals (Fig. 3). This finding is interesting, as it has been proposed that the contamination of the environment with such metals may play a role in the maintenance of AMR genes in the microbial communities in the soil (Alonso et al., 2001; Baker-Austin et al., 2006). In another recent study, Di Cesare and colleagues also reported the co-occurrence of heavy metals and antibiotic resistance genes within the resident bacterial communities in urban WWTPs, in the different steps of the treatment process. Moreover, these authors could successfully isolate living microorganisms carrying such resistance genes (Di Cesare et al., 2016). These observations confirm that the sludges in Italy may be a source of antimicrobial resistance genes and that their release in the environment through the land application of TSIs may pose a risk for the diffusion of AMR to the soil and the crop bacterial communities.

The taxonomic profiling of the TSI highlighted differences in the composition of the microbiota between the analysed samples composed of sludges, compost and mixed compost, showing a higher richness in taxonomic units in the former sample types. This finding is not surprising as the sludges samples were from wastewater treatment plants collecting inputs from urban and livestock wastes.

A deeper analysis showed that some classes of the Actinobacteria phylum prevailed in the compost samples (CO1 and CO2, and to a lesser extent MCO and CO3), while showed a much lower representation in the four analysed sludge-based BSO. This phylum shows a wide diffusion in the aquatic and terrestrial ecosystem (Alvarez et al., 2016) and its abundance in the compost-based samples analysed here can be explained by the inclusion of the matter from the pruning of the trees of

the urban areas in the composting process. Similarly, in the Eukaryota domain, the three compost samples exhibited a high content of Ascomycota phylum, which appeared reduced in the MCO and even lower in all the sludges (Fig. 1). Beside the implication of the association between specific operational taxonomic units (OTU) and the source of the samples, the taxonomy profiles of the TSIs obtained in this study allowed us to group the samples based on their origin. The PCoA analysis done using the data on the taxonomy confirmed the ability of the metagenomics profiling to identify separate groups of samples based on their taxonomic composition (Fig. 2B). In particular, the sludge-based BSO samples all grouped on the right of the chart, while the compost-based CO samples placed on the left. The mixed compost MCO sample was represented by the closest circle to the line separating the two groups (Fig. 2B). This topology reflects the declared presence of sludge in the mixed compost, which is allowed to contain up to 30% of this matter according to the Italian legislation. Similarly, the analysis performed using the COMMET software also showed a net separation between the group of compost and the sludge samples (Fig. 2A). This observation is interesting as the COMMET tool performs a comparison between the sequence content of the metagenomes and does not perform any annotation and OTU assignment. This analysis could allow identifying the origin of a TSI in a quick and easy way and could enable the risk evaluators to promptly respond to regulatory requirements.

The results of this study indicate that metagenomics is efficient to detect genomic traits associated with pathogens and antimicrobial resistance when comprehensive databases are available. Moreover, the taxonomic profiling of TSIs could allow to identify the source of TSI samples with unknown origin, making metagenomics a promising tool for the traceability of TSI in presence of a regulatory framework establishing limits in terms of presence of sludges in compost used to amend agricultural soils. This could be particularly useful to improve the development of a harmonized legislation stating that top soil improvers not falling within the scope of the End of Waste criteria can't be used as fertilizers, as considered at highest risk for human and animal health.

Funding sources supporting the paper

The Italian Ministry of Health grant no. RF-2009-1534860, "ENVI-FOOD" project, is acknowledged.

Appendix A. Supplementary material

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.envres.2017.02.004.

References

10/07/2013 (Dm) Politiche agricole. Aggiornamento degli allegati del Dlgs 29 aprile 2010, n. 75, concernente il riordino e la revisione della disciplina in materia di fertilizzanti. Ministero delle politiche agricole alimentari e forestali.

2008/98/EC (Dir.) Directive of the European Parliament and of the Council of 19 November 2008 on Waste and Repealing Certain Directives. EU(ed.).

86/278/EEC (Dir.) Council Directive 86/278/EEC of 12 June 1986 on the Protection of the Environment, and in Particular of the Soil, When Sewage Sludge is used in Agricolture. EU(ed.).

99/1992 (D.lgs) Attuazione della direttiva 86/278/CEE concernente la protezione dell'ambiente, in particolare del suolo, nell'utilizzazione dei fanghi di depurazione in agricoltura.

Alonso, A., Sánchez, P., Martínez, J.L., 2001. Environmental selection of antibiotic resistance genes. Environ. Microbiol. 3, 1–9.

Alvarez, A., Saez, J.M., Davila Costa, J.S., Colin, V.L., Fuentes, M.S., Cuozzo, S.A., et al., 2016. Actinobacteria: current research and perspectives for bioremediation of pesticides and heavy metals. Chemosphere 166, 41–62.

Baker-Austin, C., Wright, M.S., Stepanauskas, R., McArthur, J.V., 2006. Co-selection of antibiotic and metal resistance. Trends Microbiol. 14, 176–182.

Bengtsson-Palme, J., Hammarén, R., Pal, C., Östman, M., Björlenius, B., Flach, C.F., et al., 2016. Elucidating selection processes for antibiotic resistance in sewage treatment plants using metagenomics. Sci. Total Environ..

Binh, C.T., Heuer, H., Kaupenjohann, M., Smalla, K., 2008. Piggery manure used for soil fertilization is a reservoir for transferable antibiotic resistance plasmids. FEMS Microbiol. Ecol. 66, 25–37.

Brambilla, G., Abate, V., De Filippis, S.P., Fulgenzi, A.R., Iamiceli, A.L., Mazzette, A., et al., 2011. Polychlorodibenzodioxin and -furan (PCDD and PCDF) and dioxin-like polychlorobiphenyl (DL-PCB) congener levels in milk of grazing sheep as indicators of the environmental quality of rural areas. J. Agric Food Chem. 59, 8513–8517.

Brambilla, G., Abate, V., Battacone, G., De Filippis, S.P., Esposito, M., Esposito, V., Miniero, R., 2016. Potential impact on food safety and food security from persistent organic pollutants in top soil improvers on Mediterranean pasture. Sci. Total Environ. 543, 581–590.

Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60.

Di Cesare, A., Eckert, E.M., D'Urso, S., Bertoni, R., Gillan, D.C., Wattiez, R., Corno, G., 2016. Co-occurrence of integrase 1, antibiotic and heavy metal resistance genes in municipal wastewater treatment plants. Water Res. 94, 208–214.

Gotelli, J.N., Colwell, K.R., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol. Lett. 4, 379–391.

Gottschall, N., Topp, E., Metcalfe, C., Edwards, M., Payne, M., Kleywegt, S., et al., 2012. Pharmaceutical and personal care products in groundwater, subsurface drainage, soil, and wheat grain, following a high single application of municipal biosolids to a field. Chemosphere 87, 194–203.

Hao, H., Dai, M., Wang, Y., Huang, L., Yuan, Z., 2012. Key genetic elements and regulation systems in methicillin-resistant Staphylococcus aureus. Future Microbiol. 7, 1315–1329.

Helmi, K., Skraber, S., Gantzer, C., Willame, R., Hoffmann, L., Cauchie, H.M., 2008. Interactions of Cryptosporidium parvum, Giardia lamblia, vaccinal poliovirus type 1, and bacteriophages phiX174 and MS2 with a drinking water biofilm and a wastewater biofilm. Appl. Environ. Microbiol. 74, 2079–2088.

Heuer, H., Smalla, K., 2007. Horizontal gene transfer between bacteria. Environ. Biosaf. Res. 6, 3–13.

Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data. Genome Res. 17, 377–386.

ISPRA, 2012. Istituto Superiore per la Protezione e la Ricerca Ambientale. Rapporto Rifiuti Urbani.

ISWA, 2013. Working Group on Biological Treatment of Waste. FACT SHEET: Requirements for the Treatment and Application of Digestate and Sewage Sludge (Biosolids) - Cases of several countries. International Solid Waste Association.

Joensen, K.G., Scheutz, F., Lund, O., Hasman, H., Kaas, R.S., Nielsen, E.M., Aarestrup, F.M., 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. J. Clin. Microbiol. 52, 1501–1510.

Kanehisa, M., Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Liang, C., Das, K.C., McClendon, R.W., 2003. The influence of temperature and moisture contents regimes on the aerobic microbial activity of a biosolids composting blend. Bioresour. Technol. 86, 131–137.

López-Gálvez, F., Truchado, P., Sánchez, G., Aznar, R., Gil, M.I., Allende, A., 2016. Occurrence of enteric viruses in reclaimed and surface irrigation water: relationship with microbiological and physicochemical indicators. J. Appl. Microbiol. 121, 1180–1188.

Maillet, N., Collet, G., Vannier, T., Lavenier, D., Peterlongo, P., 2014. COMMET: comparing and Combining Multiple Metagenomic Datasets. IEEE BIBM. Belfast, United Kingdom. hal-01080050.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., et al., 2008. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinform. 9, 386.

Mezzanotte, V., Antonelli, M., Citterio, S., Nurizzo, C., 2007. Wastewater disinfection alternatives: chlorine, ozone, peracetic acid, and UV light. Water Environ. Res. 79, 2373–2379.

Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.Y., Cohoon, M., et al., 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res. 33, 5691–5702.

Piddock, L.J., 2006. Multidrug-resistance efflux pumps - not just for resistance. Nat. Rev. Microbiol. 4, 629–636.

Pruitt, K.D., Tatusova, T., Maglott, D.R., 2005. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res. 33, D501–D504.

Saveyn, H., Eder, P., 2014. End-of-Waste Criteria for Biodegradable Waste Subjected to Biological Treatment (compost & digestate): Technical proposals. EU Commission, IPTS Sevilla (Spain).

Sengeløv, G., Agersø, Y., Halling-Sørensen, B., Baloda, S.B., Andersen, J.S., Jensen, L.B., 2003. Bacterial antibiotic resistance levels in Danish farmland as a result of treatment with pig manure slurry. Environ. Int, 28, 587–595.

Smith, G.M., Zuur, A.F., Ieno, E.N., 2007. Analysing Ecological Data. Springer, New york, 672.

Suominen, K., Verta, M., Marttinen, S., 2014. Hazardous organic compounds in biogas plant end products–soil burden and risk to food safety. Sci. Total Environ. 491–492, 192–199.

Tatusov, R.L., Koonin, E.V., Lipman, D.J., 1997. A genomic perspective on protein families. Science 278, 631–637.

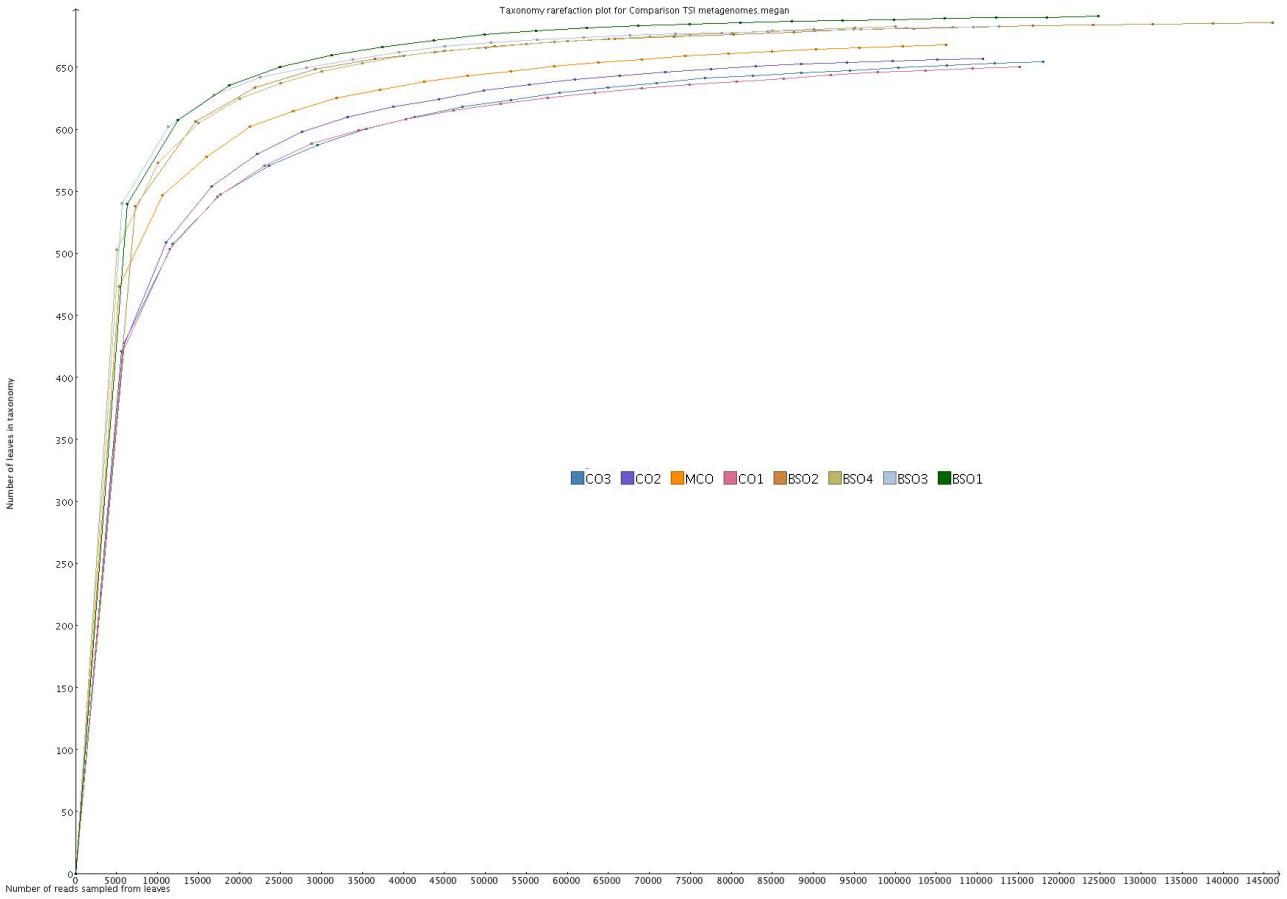
Tozzoli, R., Di Bartolo, I., Gigliucci, F., Brambilla, G., Monini, M., Vignolo, E., et al., 2016. Pathogenic Escherichia coli and enteric viruses in biosolids and related top soil improvers in Italy. J. Appl. Microbiol..

Trautmann, N., Olynciw, E., 2010. Composting-Compost Microorganisms. (WWW

document). URL.
Usman, K., Khan, S., Ghulam, S., Khan, M., Khan, N., Khan, M., Khalil, S., 2012. Sewage sludge: an important biological resource for sustainable agriculture and its environmental implications. Am. J. Plant Sci. 3 (12), 1708–1721.
van Hoek, A.H., Aarts, H.J., Bouw, E., van Overbeek, W.M., 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.
Yergeau, E., Masson, L., Elias, M., Xiang, S., Madey, E., Huang, H., et al., 2016. Comparison of methods to identify pathogens and associated virulence functional genes in biosolids from two different wastewater treatment facilities in Canada. PLoS One 11, e0153554.
Zennegg, M., Munoz, M., Schmid, P., Gerecke, A.C., 2013. Temporal trends of persistent organic pollutants in digested sewage sludge (1993–2012). Environ. Int. 60, 202–208.
Zhou, N., Lv, D., Wang, S., Lin, X., Bi, Z., Wang, H., et al., 2016. Continuous detection and genetic diversity of human rotavirus A in sewage in eastern China, 2013–2014. Virol. J. 13, 153.

Supplementary Material

Figure S1 Rarefaction curves of the studied metagenomes



Chapter 6

Metagenomic characterization of the human intestinal microbiota in faecal samples from STEC-infected patients

6. Metagenomic characterization of the human intestinal microbiota in faecal samples from STEC-infected patients

The metagenomics proved to be an efficient tool to perform a direct and complete characterization of microbial communities in complex matrices (Chapter 5). In this Chapter we developed an enhanced metagenomic approach to investigate the relationships between STEC and the human intestinal microbiota, once they establish in the human intestine upon infection. We examined changes in the microbial composition of the intestinal microbiota in faecal samples from patients suffering from STEC infection compared to healthy and healed controls, collected during an outbreak caused by a STEC O26:H11. In this study, we have faced the metagenomic analysis by adopting two different bioinformatic approaches, one based on the mapping of the reads to databases in order to achieve a deep taxonomic classification of the stool samples, and the other based on the reconstruction, from the metagenomes, of putative entire genomes followed by the annotation and classification of the bins obtained into taxonomic units. Both the analyses returned the same picture of a different microbial distribution in STEC positive samples in comparison to those negative for the presence of the infection. In general, faeces collected from infected patients showed a lower intestinal abundance of the beneficial microorganisms *Bifidobacterium* and *Clostridiales* spp. with respect to those from the controls group, where those microorganisms predominated. This study provides the first evidence of the changes occurring in the human intestinal microbiota in the course of STEC infection. This observation parallels what it has been observed in other studies dealing different infections or intestinal afflictions, suggesting a possible common response of the microbiota to intestinal disorders. These findings suggested the occurrence of a possible competition between beneficial microorganisms and STEC strains in human intestine and opened the way to the study of measures to mitigate the impact of STEC infections and their progression towards the most severe forms.

Additionally, the metagenomic sequencing allowed to identify in the STEC positive samples, all the virulence traits present in the genomes of the STEC O26 that caused the outbreak, suggesting that metagenomics may be a promising tool for the culture-independent clinical diagnosis of the infections.

frontiers in Cellular and Infection Microbiology

ORIGINAL RESEARCH
published: 06 February 2018
doi: 10.3389/fcimb.2018.00025

Metagenomic Characterization of the Human Intestinal Microbiota in Fecal Samples from STEC-Infected Patients

Federica Gigliucci[1,2], F. A. Bastiaan von Meijenfeldt[3], Arnold Knijn[1], Valeria Michelacci[1], Gaia Scavia[1], Fabio Minelli[1], Bas E. Dutilh[3,4], Hamideh M. Ahmad[5], Gerwin C. Raangs[5], Alex W. Friedrich[5], John W. A. Rossen[5] and Stefano Morabito[1]*

[1] Department of Food Safety, Nutrition and Veterinary Public Health, Istituto Superiore di Sanità, Viale Regina Elena, Rome, Italy, [2] Department of Sciences, University Roma Tre, Rome, Italy, [3] Theoretical Biology and Bioinformatics, Utrecht University, Utrecht, Netherlands, [4] Centre for Molecular and Biomolecular Informatics, Radboud University Medical Centre, Nijmegen, Netherlands, [5] Department of Medical Microbiology, University of Groningen, University Medical Center Groningen, Groningen, Netherlands

The human intestinal microbiota is a homeostatic ecosystem with a remarkable impact on human health and the disruption of this equilibrium leads to an increased susceptibility to infection by numerous pathogens. In this study, we used shotgun metagenomic sequencing and two different bioinformatic approaches, based on mapping of the reads onto databases and on the reconstruction of putative draft genomes, to investigate possible changes in the composition of the intestinal microbiota in samples from patients with Shiga Toxin-producing *E. coli* (STEC) infection compared to healthy and healed controls, collected during an outbreak caused by a STEC O26:H11 infection. Both the bioinformatic procedures used, produced similar result with a good resolution of the taxonomic profiles of the specimens. The stool samples collected from the STEC infected patients showed a lower abundance of the members of *Bifidobacteriales* and *Clostridiales* orders in comparison to controls where those microorganisms predominated. These differences seemed to correlate with the STEC infection although a flexion in the relative abundance of the Bifidobacterium genus, part of the *Bifidobacteriales* order, was observed also in samples from Crohn's disease patients, displaying a STEC-unrelated dysbiosis. The metagenomics also allowed to identify in the STEC positive samples, all the virulence traits present in the genomes of the STEC O26 that caused the outbreak as assessed through isolation of the epidemic strain and whole genome sequencing. The results shown represent a first evidence of the changes occurring in the intestinal microbiota of children in the course of STEC infection and indicate that metagenomics may be a promising tool for the culture-independent clinical diagnosis of the infection.

Keywords: STEC, microbiota, human gut, HUS, diarrhea, metagenomics

OPEN ACCESS

Edited by:
Alfredo G. Torres, University of Texas Medical Branch, United States

Reviewed by:
Maite Muniesa, University of Barcelona, Spain
Analía Inés Etcheverría, National University of Central Buenos Aires, Argentina
Wessam Galia, UMR5557 Ecologie Microbienne (LEM), France

****Correspondence:***
Federica Gigliucci
federica.gigliucci@libero.it

Received: *30 October 2017*
Accepted: *18 January 2018*
Published: *06 February 2018*

Citation:
Gigliucci F, von Meijenfeldt FAB, Knijn A, Michelacci V, Scavia G, Minelli F, Dutilh BE, Ahmad HM, Raangs GC, Friedrich AW, Rossen JWA and Morabito S (2018) Metagenomic Characterization of the Human Intestinal Microbiota in Fecal Samples from STEC-Infected Patients. Front. Cell. Infect. Microbiol. 8:25. doi: 10.3389/fcimb.2018.00025

Abbreviations: STEC, Shiga toxin-producing *Escherichia coli;* HUS, Haemolytic Uremic Syndrome; CD, Crohn's Disease; EPEC, Enteropathogenic *Escherichia coli*; OTUs, Operational Taxonomic Units.

INTRODUCTION

Shiga toxin-producing *Escherichia coli* (STEC) are a heterogeneous *E. coli* pathogroup causing food-borne outbreaks and sporadic cases of disease worldwide (Armstrong et al., 1996). STEC may cause severe afflictions in humans due to their ability to produce potent cytotoxins, the Shiga toxins (Stx), acquired upon infection with bacteriophages carrying stx genes, which can remain stably integrated into the bacterial chromosome (O'Brien et al., 1984). Stx exert their action by blocking the protein synthesis in the target cells by inactivating ribosomes (Okuda et al., 2006). Upon infection, the host can present a wide range of symptoms, including uncomplicated diarrhea, haemorrhagic colitis and the life-threatening haemolytic uremic syndrome (HUS).

The pathogenesis of STEC infections is not completely understood as it seems that, beside the virulence potential of the infecting strains, a number of other factors appear to be involved in the progression of the clinical symptoms. One possibility is that the human intestinal microbiota play a role by interfering with the ability of STEC to efficiently colonize the gastro-intestinal tract, as it has been proposed for other bacterial infections (Fujiwara et al., 2001; Gueimonde et al., 2007). Additionally, Gamage and colleagues proposed that different bacterial species in the host microbiota can act as amplifiers of the Stx-converting phage resulting in an augmented ability to produce the toxin (Gamage et al., 2003).

More and more data are becoming available on the role of human microbiota in health and disease and there is increasing evidence that the commensal bacteria play a crucial role in protecting human health (Forbes et al., 2016). Indeed, the human gut microbiota is a homeostatic ecosystem with several vital functions essential to host health, including protection against pathogens (Shreiner et al., 2015). Tap and colleagues showed that the human gut microbiota is governed by the presence of *Bacteroidetes*, *Firmicutes*, *Actinobacteria*, *Proteobacteria* and, in some cases *Verrucomicrobia* bacterial phyla (Tap et al., 2009). Members of these taxa, including *Faecalibacterium*, *Ruminococcus*, *Eubacterium*, *Dorea*, *Bacteroides*, *Alistipes,* and *Bifidobacterium* genera, constitute a phylogenetic core shared among individuals (Tap et al., 2009). In particular, it has been shown that species belonging to *Bifidobacterium* genus and butyrate-producing bacteria, belonging to the *Clostridiales* order, might exert a variety of beneficial health effects (O'Callaghan and van Sinderen, 2016; Rivière et al., 2016). Hence, a decrease in the relative abundances of *Bifidobacterium* species in the human colon has been associated with several disorders, such as inflammatory bowel disease, Crohn's disease and ulcerative colitis, irritable bowel syndrome, colorectal cancer, and increased gut permeability (O'Callaghan and van Sinderen, 2016). The mentioned disorders lead to a general change in the gut microbiota composition, favoring also the colonization and proliferation of pathogenic microorganisms (de Vos and de Vos, 2012).

We used a shotgun metagenomic sequencing approach to investigate the taxonomic composition of the gut microbiota in fecal samples taken from patients suffering from STEC infection and compared the results with those found in samples from healthy controls. Additionally, we analyzed fecal samples from patients with Crohn's disease with and without evidence of infection with STEC and added the microbial profiles of these samples to the analysis in order to assess if any change observed in the composition of the gut microbiota in course of STEC infections may be related with a dysbiotic status rather than associated with the STEC infection itself.

MATERIALS AND METHODS

Samples Origin

Fecal samples ($N = 10$) from children (ages 0–4) were collected during an outbreak of STEC infections in 2015 in a nursery in the province of Rome, Italy. During the investigation, a STEC O26:H11 was isolated from different patients. Three samples were from patients with diarrhea (A.9, A.8, A.14), three were collected from patients after 2 weeks from the restoration of the normal intestinal function (A.40, A.32, A.41), and four samples (A.4, A.30, A.16, 481-5) from healthy subjects. In addition, four stool samples from four patients (ages 10–18) suffering from Crohn's Disease (CD) and hospitalized at the University Medical Center Groningen, The Netherlands, were included in the analysis (Samples 1, 2, 5, and 6). These samples were previously analyzed by real-time PCR for the presence of virulence genes associated with STEC (*stx1* and *stx2*) and EPEC (*escV*) (Gauthier et al., 2003) pathogroups, revealing the presence of *stx2* gene in Sample 2, the presence of the *escV* gene in Sample 1 and 5, and none of the mentioned virulence genes in Sample 6. The CD samples used for the present analyses were collected in the course of routine diagnostics and infection prevention controls. Oral consent for the use of such clinical samples for research purposes is routinely obtained upon patient admission to the UMCG, in accordance with the guidelines of the Medical Ethics Committee of the University Medical Center Groningen. The experiments, accordingly with the guidelines of the Declaration of Helsinki and the institutional regulations, were performed on anonymized samples.

DNA Extraction and Sequencing

DNA of the specimens related with the Italian outbreak was extracted from 0.20 g of each stool sample using the EZNA Stool DNA extraction kit (Omega Bio-tek, Norcross, Ga.) following manufacturer's instructions, whereas for the CD samples, DNA was extracted from 0.25 g feces using the Power Soil DNA isolation kit (MO BIO Laboratories inc., Carlsbad, CA, USA) following the manufacturer's instructions. We did not observe marked differences in the purity of DNA produced with the two methods as assessed by considering the ratio between the absorbance measured at 260 and 280 nm.

Sequencing libraries were prepared from 100 ng of the DNA extracted from samples **A.9, A.4, and A.30**, using the NEBNext Fast DNA Fragmentation & Library Prep kit (New England BioLabs, New England, USA). In detail, the DNA was enzymatically fragmented to obtain fragments of about 400 bp, through an incubation at 25°C for 15 min, followed by 10 min at 70°C. The fragmented DNA was subjected to link with adapters

and size selection of 450 bp fragments by electrophoresis on E-Gel SizeSelect 2% (Invitrogen, Carlsbad, USA) followed by PCR amplification as indicated in the NEBNext Fast DNA Fragmentation & Library Prep kit manual (New England BioLabs, USA). The libraries were amplified individually through emulsion PCR with an Ion OneTouch 2 and sequenced with an Ion Torrent Personal Genome Machine (Life Technologies, 118 Carlsbad, USA), using the 400 bp sequencing protocol. The three samples were sequenced individually in three different runs using a 316 V2 chip per run.

Sequencing of the samples **A.8, A.14, A.40, A.32, A.41, A.16, 481-5** as well as of the four CD fecal samples DNA was carried out using the TruSeq Nano DNA Library Preparation kit (Illumina, San Diego, CA, US) and a MiSeq platform (Illumina, San Diego, CA, US). In detail, sequencing libraries were prepared from 100 ng of the DNA extracted from each sample. DNA was mechanically fragmented to obtain a 350 bp insert size, using the M220 Focused-ultrasonicator™ based on Covaris AFA (Adaptive Focused Acoustics™) technology. The fragmented DNA was subjected to end repair and size selection of 350 bp fragments, followed by adenylation of 3' ends, link with adaptors and a final enrichment of DNA fragments, following the TruSeq Nano DNA Library Preparation kit manual (Illumina, San Diego, CA, US). The 11 samples were sequenced in three different runs, using a 600V3 cartridge per run generating 300 bp paired-end reads.

All the metagenomes are available at European Nucleotide Archive at EMBL-EBI under the accession number PRJEB23207.

Bioinformatic Analyses

Read Mapping Analysis

Reference-based bioinformatic analyses of the metagenomes were performed using the tools available on the ARIES public webserver (https://www.iss.it/site/aries/).

In detail, raw sequence reads were subjected to a quality check and trimmed to remove the adaptors and to accept 20 as the lowest Phred value. The identification of the presence of *E. coli* virulence genes was performed through the pipeline Virulotyper, which employs the Bowtie2 v2.3.2.2 program (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) (Langmead and Salzberg, 2012) to map the sequencing reads against the *E. coli* Virulence genes database (Joensen et al., 2014). Virulence genes showing average coverage above 1X were considered to be present in the sample.

Taxonomic analysis was performed using the DIAMOND v0.8.24 tool (Buchfink et al., 2015) to align the reads in FASTA format to the NCBI-nr (non-redundant) database (ftp://ftp.ncbi.nlm.nih.gov/) downloaded on July 2016, with the DIAMOND-BLASTX algorithm. Visualization was done with MEGAN (MEta Genome ANalyzer) version 6 (http://www-ab.informatik.uni-tuebingen.de/software/megan6/) (Huson et al., 2007).

The operational taxonomic units (OTUs) content of the samples was determined using MEGAN 6 and the converter script prot-gi2taxid-August2016X.bin downloaded from the NCBI website (https://www.ncbi.nlm.nih.gov). MEGAN 6 was also used to perform rarefaction analysis (Gotelli and Colwell, 2001).

Reference Free Analysis

The same metagenomes were also analyzed using a novel approach, as described by Kang et al. (2015), which allows assessing the degree of similarity between complex microbial communities through the reconstruction of draft genome sequences from shotgun metagenomic sequencing. Reads from the different metagenomes obtained with the Illumina platform were de novo assembled. The assembled genomic fragments were eventually grouped in putative genomes (bins) using probabilistic distances. For a visualization of the binning process we refer to Figure 1 in Kang et al. (2015). Default settings were used for all programs, unless otherwise mentioned.

In detail, the paired-end trimmed Illumina reads were cross-assembled with SPAdes v3.10.1 (http://bioinf.spbau.ru/spades) (Bankevich et al., 2012) in "—meta" mode (Nurk et al., 2017) and used as reference for mapping of all the single metagenomes, including those obtained with the Ion Torrent PGM. The quality of the cross-assembled scaffolds was evaluated using the QUality ASsessment Tool (QUAST) v4.5 (http://quast.sourceforge.net/) (Gurevich et al., 2013). QUAST provides assembly statistics including the number of assembled contigs, the length of the longest contig and the values of N50, N75, L50, L75, GC (%).

The reads of the metagenomes were mapped to the cross-assembled scaffolds using Burrows-Wheeler Aligner (BWA) v0.7.15, with the BWA-MEM algorithm (Li and Durbin, 2010). BWA-MEM was run with default settings, thus distributing reads that map to multiple places evenly, as suggested in the MetaBAT usage manual (https://bitbucket.org/berkeleylab/metabat). The output of BWA-MEM was converted from SAM format to the binary BAM format with the tools provided in the SAMtools suite (Li et al., 2009). The percentage of the reads mapping against the cross-assembly was estimated for the different metagenomes using SAMtools flagstat. The open source software MetaBAT2 (Metagenome Binning with Abundance and Tetra-nucleotide frequencies) v2.9.1 (http://bitbucket.org/berkeleylab/metabat) (Kang et al., 2015) was used to obtain bins from the scaffolds. MetaBAT2 was fed with the cross-assembly as input together with a depth file based on the bam files. The depth file was generated with the script jgi_summarize_bam_contig_depths, that is supplied with MetaBAT. The quality of the genomic bins generated was assessed with CheckM v1.0.5 (Parks et al., 2015) in "—lineage_wf" mode. CheckM assesses bin completeness, contamination, and strain heterogeneity based on the absence and presence of sets of expected single copy marker genes. Bins with a completeness >35% and a contamination <5% were selected for further analyses.

The scaffolds were annotated with the CAT (The Contig Annotation Tool) pipeline (Cambuy et al., 2016) and used to taxonomically classify the bins. To account for possible conflicts within a bin, we only annotated a bin to a taxonomic level if at least half of the length of the sequences within the bin showed a consistent annotation. If no annotation reached majority, we annotated the bin to the annotation with the longest length representation, but marked that annotation as "possible." For instance, if 70% of the length of the bin is annotated on the genus level to *Escherichia*,

but only 30% is annotated to *E. coli* on the species level, that bin is annotated as an *Escherichia* bacterium, possibly *E. coli*.

To calculate relative abundance of the bins in the samples, we used the depth file generated earlier with jgi_summarize_bam_contig_depths. Average coverage of a bin in a sample was calculated by multiplying the depth of the scaffolds in that bin by their length, and dividing their sum by the total base pair length of all the scaffolds in the bin. Abundance was made relative per sample by dividing this average read coverage per base pair by the sum of average read coverage per base pair for all the bins. Bin abundance was normalized to account for the fact that not all reads mapped to bins, by multiplying relative abundance with the fraction of binned reads, which was calculated as the sum of reads mapping to binned scaffolds (as calculated with samtools idxstats) divided by the total number of reads mapping to the cross-assembly (as calculated with samtools flagstat). Thus, the sum of relative abundances of all bins adds up to the fraction of total reads mapping to the cross-assembly that map to the bins in a sample.

The scaffolds were searched for the occurrence of the genes present in the *E. coli* Virulence genes database (Joensen et al., 2014). The database was queried with tblastx v2.6.0+ (Camacho et al., 2009) against the scaffolds. A gene was considered present on a scaffold if its hit had an e-value below 1e-5 and query coverage of at least 70%.

RESULTS

Detection of Virulence Genes Associated with Pathogenic *E. coli* in the Metagenomes

Detection of the STEC O26:H11 Virulence Genes

The virulence features of the STEC O26:H11 strain that had caused the outbreak had been previously identified by characterizing the isolated strain through real-time PCR and Whole Genome Sequencing (unpublished). In this work, we first checked the capability of metagenomics to identify the presence of the genes associated with the epidemic STEC strain. Mapping of the quality-controlled reads against the *E. coli* Virulence genes database (Joensen et al., 2014) showed the presence of all the genes composing the virulome of the STEC O26 outbreak strain (**Table 1**). In one sample (A.9), the metagenomics did not show the presence of the *espF* gene, while in sample A.14 we identified the presence of the gene *iroN*, which however was not present in the outbreak strain (**Table 1**). This gene was later identified in a scaffold that was annotated to the family *Enterobacteriaceae*, but was unbinned.

TABLE 1 | Comparison of the STEC virulence genes identified in the samples collected from patients with STEC infection by metagenomics with those obtained through WGS of the isolated outbreak strain.

Metagenomic analysis										Isolation based method
Patients with diarrhea			Recovered patients			Healthy subjects				
A. 9	A. 8	A. 14	A. 40	A. 41	A. 32	A. 4	A. 16	481-5	A. 30	Epidemic strain
Cif	*cif*	*cif*	***ehxA***	*mchF*	***katP***	*epeA*	*cba*		*senB*	*cif*
Eae	*eae*	*eae*	***espP***	*tsh*	*lpfA*	***espP***	*cma*			*eae*
efa1	*efa1*	*efa1*	***katP***			*gad*	*gad*			*efa1*
ehxA	***ehxA***	***ehxA***	*mchC*			*iha*	*ireA*			***ehxA***
espA	*espA*	*espA*	*pic*				*iss*			*espA*
espB	*espB*	*espB*	*senB*				*lpfA*			*espB*
espJ	*espF*	*espF*	***toxB***				*pic*			*espF*
espP	*espJ*	*espJ*					*vat*			*espJ*
gad	***espP***	***espP***								***espP***
iha	*gad*	*gad*								*gad*
iss	*iha*	*iha*								*iha*
katP	*iss*	*iroN*								*iss*
lpfA	***katP***	*iss*								***katP***
nleA	*lpfA*	***katP***								*lpfA*
nleB	*nleA*	*lpfA*								*nleA*
nleC	*nleB*	*nleA*								*nleB*
stx2a	*nleC*	*nleB*								*nleC*
tir	*stx2a*	*nleC*								*stx2a*
toxB	*tir*	*stx2a*								*tir*
	toxB	*tir*								***toxB***
		toxB								

Plasmidic genes normally present on the large virulence plasmid of STEC O26 are indicated in bold.

The genes encoding the Stx were not identified in the specimens collected from healthy subjects and from recovered patients. However, we could observe in some of the latter metagenomes the presence of genes carried by the large virulence plasmid of STEC and identified also in the outbreak strain (**Table 1**). In particular, the metagenome from sample A.40 displayed the presence of plasmidic genes *ehxA* (Beutin et al., 1990), *espP* (Brunder et al., 1997), *katP* (Caprioli et al., 2005), and *toxB* (Tatsuno et al., 2000), while that from sample A.32 had the gene *katP* (Caprioli et al., 2005) (**Table 1**). Finally, we could not observe the presence of any gene associated with STEC in any of the metagenomes from healthy subjects, with the exception of the presence of *espP* in sample A.4 and *lpfA* in sample A.16 (**Table 1**).

The same analysis carried out on the metagenomes of the samples collected from patients with Crohn's disease confirmed the previous real-time PCR results (**Table 2**). In detail, Sample 6 appeared negative for the presence of virulence genes associated with pathogenic *E. coli* infections; Sample 2 showed the presence of the Stx2f encoding genes, confirming the previous evidence of STEC infection; finally, the remaining two samples, although they did not show the presence of the *escV* gene, previously identified by real-time PCR, displayed a virulence genes profile compatible with the presence of an aEPEC strain (**Table 2**).

Detection of Virulence Genes in the Cross-Assembled Scaffolds

In a complementary bioinformatic analysis of the same data, we assembled the metagenomic reads into scaffolds, and binned the scaffolds into draft genome sequences (Garza and Dutilh, 2015). The cross-assembly of the Illumina metagenomes, including the samples from STEC infections and the related control group as well as those from Crohn's disease, produced 429,862 scaffolds of size $\geq$500 bp, where the longest scaffold was 505,337 bp (Table S1). The percentage of the reads mapping against the assembled scaffolds was comparable between the different metagenomes, and seemed only slightly influenced by sequencing platform, ranging from 85.96 to 99.77% (Table S2). It is important to note that the cross-assembly is only based on the Illumina sequences, but the Ion Torrent reads showed high mappability values as well (Table S2). Finally, the metagenome binning produced 209 draft genome bins in total.

The tblastx search confirmed the presence of *E. coli* virulence genes in the cross-assembled metagenomic scaffolds, that were already detected in the metagenomes based on mapping the quality-controlled reads on the virulence genes database as described above. Additionally, this analysis allowed to localize each gene in specific scaffolds and within binned draft genomes. Most of the assembled scaffolds that contained virulence genes

TABLE 2 | Comparison of virulence genes associated with pathogenic *E. coli* infection identified by metagenomics and real-time PCR in the samples collected from patients with Crohn's disease.

Sample 1		Sample 2			Sample 5		Sample 6	
Metagenomics reads mapping	**Real time PCR**	**Metagenomics reads mapping**	**Metagenomics bin alignment**	**Real time PCR**	**Metagenomics reads mapping**	**Real time PCR**	**Metagenomics reads mapping**	**Real time PCR**
cba	*escV*	*cif*	*eae*	*stx2f*	*cif*	*escV*		
cif		*cnf1*	*gad*		*espA*			
cma		*eae*	*hlyE*		*espF*			
eae		*espA*	*iha*		*gad*			
espA		*espC*	*ireA*		*nleB*			
espJ		*espF*	*iroN*					
espD		*espJ*	*katP*					
gad		*fim41a*	*lpfA*					
iroN		*gad*	*mchF*					
iss		*iha*	*pet*					
katp		*ireA*	*prfB*					
lpfA		*iroN*	*sat*					
mchF		*iss*						
pet		*lpfA*						
tir		*mchB*						
tsh		*mchC*						
vat		*mchF*						
		mcmA						
		nleB						
		nleC						
		pic						
		stx2f						
		tir						
		vat						

that were present in the epidemic strain were unbinned, but were annotated as *E. coli* on the scaffold level. Moreover, all of the plasmidic genes that were seen in recovered patients with read mapping were also unbinned. It is common that plasmids are not associated with draft genomes binned from metagenomes, as their genomic signals including abundance and nucleotide usage are different from those of the core genome (Beitel et al., 2014; Kang et al., 2015). Not all genes seen with read mapping were found in the scaffolds, the most notable omission being *ehxA*.

Taxonomic Profiling of the Microbiota from Stool Samples

The metagenomic sequencing produced datasets of different sizes, mainly due to the use of two different sequencing platforms. As for the samples from STEC infections and the related control group, an average of 3,563,158 reads per sample were obtained from the Ion Torrent sequencing, while an average of 12,287,432 reads per sample were obtained from the Illumina platform. Nevertheless, the rarefaction curves (Gotelli and Colwell, 2001) showed that the diversity of taxonomic units was comparable between the different samples (Figure S1). The metagenomic sequencing of the four stool samples from Crohn's disease patients, produced with the Illumina sequencer produced on average 16,429,446 reads per sample.

Taxonomic Profiling Based on Read Mapping

In the metagenomes assayed, the most abundant phyla of *Bacteria* were the *Firmicutes* (16.6–79%), *Bacteroidetes* (0.2–63.1%), *Proteobacteria* (1.65–56.6%), and *Actinobacteria* (0.8–56.2%), whereas the phylum *Verrucomicrobia* showed a high abundance only in one sample (27.4% in A.30), obtained from a healthy subject (**Table 3**). In the samples analyzed the *Bifidobacteriales* order of the *Actinobacteria* phylum appeared more abundant in the vast majority of the STEC-negative samples (**Figure 1**) and a deeper analysis showed a marked prevalence of *Bifidobacterium longum* species (Figure S2). In addition to *Actinobacteria*, also some orders of the phylum *Firmicutes* showed a different distribution in the two groups of samples (**Figure 1**) ($p < 0.05$). The *Clostridiales* (*Clostridia*) were more abundant in the control group, while the *Lactobacillales* (*Bacilli*) predominated in the cases group (**Figure 1**) ($p < 0.05$) with the exception of sample A.9, which showed a taxonomic profile more similar to that of the control group with lower *Lactobacillales* (**Figure 1**). Moreover, the *Roseburia*, *Coprococcus*, *Butyrivibrio,* and *Faecalibacterium*, previously shown to be common in the intestinal microbiota from healthy subjects (Rivière et al., 2016; Hugon et al., 2017) were the most abundant genera of *Clostridiales* in the STEC negative samples assayed in this study (**Figure 2**), except for sample A.9, which again showed a profile more similar to those obtained from the control group samples. Members of the *Proteobacteria* and *Bacteroidetes* phyla apparently were not concerned by the perturbation of the intestinal microflora following the STEC infection as their relative proportions did not show patterns attributable to specific groups in the metagenomes analyzed (**Figure 1**).

The most abundant bacterial taxa identified in the specimens collected from Crohn's disease patients were the *Bacteroidetes* (10.5–84%), *Proteobacteria* (4–77%), *Firmicutes* (9–29.5%), and *Actinobacteria* (1–26%) (**Figure 3A**). In one sample, a marked prevalence of the *E. coli* species was observed (Sample 2 in **Figure 3**). Interestingly, this sample was positive for STEC-associated genes, both in real-time PCR and at the metagenomic analysis.

Analysis of Genome Bins

For the analysis of the samples from STEC infections and the related controls, we selected 18 of the 209 bins based on values of completeness >35%, contamination <5%, and abundance >5% in at least one sample.

The abundance profile of the selected bins confirmed the results obtained with the reference-based approach, returning similar differences in the prevalence of specific taxonomic units between cases and controls (**Figure 4**).

In detail, *Enterococcus faecalis* and *Enterococcus avium* species and the *Streptococcus* genus, belonging to the order *Lactobacillales*, predominated in the STEC positive samples (**Figure 4A**). Concerning the *Clostridiales* order, members of *Peptostreptococcaceae* and *Clostridiaceae* families showed a high abundance in the cases group (**Figure 4A**), while *Lachnospiraceae* and *Ruminococcaceae* spp. (*Ruminococcus gnavus*, *Faecalibacterium prausnitzii*) showed a high representation in the set of samples from healthy and recovered subjects, confirming their association with a healthy status in the human intestine (**Figure 4B**).

Finally, the approach of reconstructing single genomes from complex microbial communities confirmed a clear prevalence of members of the *Bifidobacteriales* order in the control group (**Figure 4B**).

For the analysis of the metagenomes from Crohn's disease, we selected nine bins based on a completeness >35%, a contamination <5%, and an abundance value >2% in at least one sample.

This analysis confirmed the lower complexity of the intestinal microbiota in Crohn's disease and the high representation of *E. coli* species in Sample 2 identified with the read mapping approach (**Figure 3B**).

DISCUSSION AND CONCLUSION

The human gastrointestinal tract microflora, comprise $\sim 10^{14}$ microbial cells that live in a mutual beneficial relationship with the host (Ley et al., 2006). Indeed, the gut microorganisms have a remarkable impact on human physiology, because they modulate the normal intestinal functions, produce vitamins and contribute to obtain energy from the food (Bäckhed et al., 2005). They have also a profound influence on the local and systemic immune responses and interfere with pathogen's colonization (Shreiner et al., 2015; Forbes et al., 2016).

In the present study, we aimed at using metagenomics as a tool to investigate possible changes in the composition of the intestinal microbiota in patients with STEC infection compared to healthy controls. Shotgun metagenomic sequencing was used to identify the presence of virulence genes associated with STEC strains and to perform a taxonomic classification of the microbial communities present in fecal samples collected from subjects

TABLE 3 | Relative abundance of the most abundant bacterial phyla in the samples from STEC-infected and healthy subjects, obtained through analysis based on read mapping.

	A.9 (%)	A.8 (%)	A.14 (%)	A.40 (%)	A.41 (%)	A.32 (%)	A.4 (%)	A.16 (%)	481-5 (%)	A.30 (%)
Actinobacteria	0.5	1.1	6.3	56.2	14.8	50	14	3.5	28.5	0.8
Proteobacteria	15.7	56.6	13.3	1.9	3.4	1.6	29.5	12.2	1.6	3.7
Bacteroidetes	63.1	0.2	7.46	6.6	1	11	3.2	5	1.3	51
Firmicutes	19.7	41.3	71.2	34.2	79	36.7	52	76.4	66.5	16.7
Verrucomicrobia	0.16	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.3	27.5

Samples A.9, A.8, A.14 are from diseased patients; A.40, A.41, A.32 are from recovered patients; A.4, A.16, 481-5; A.30 are from healthy subjects.

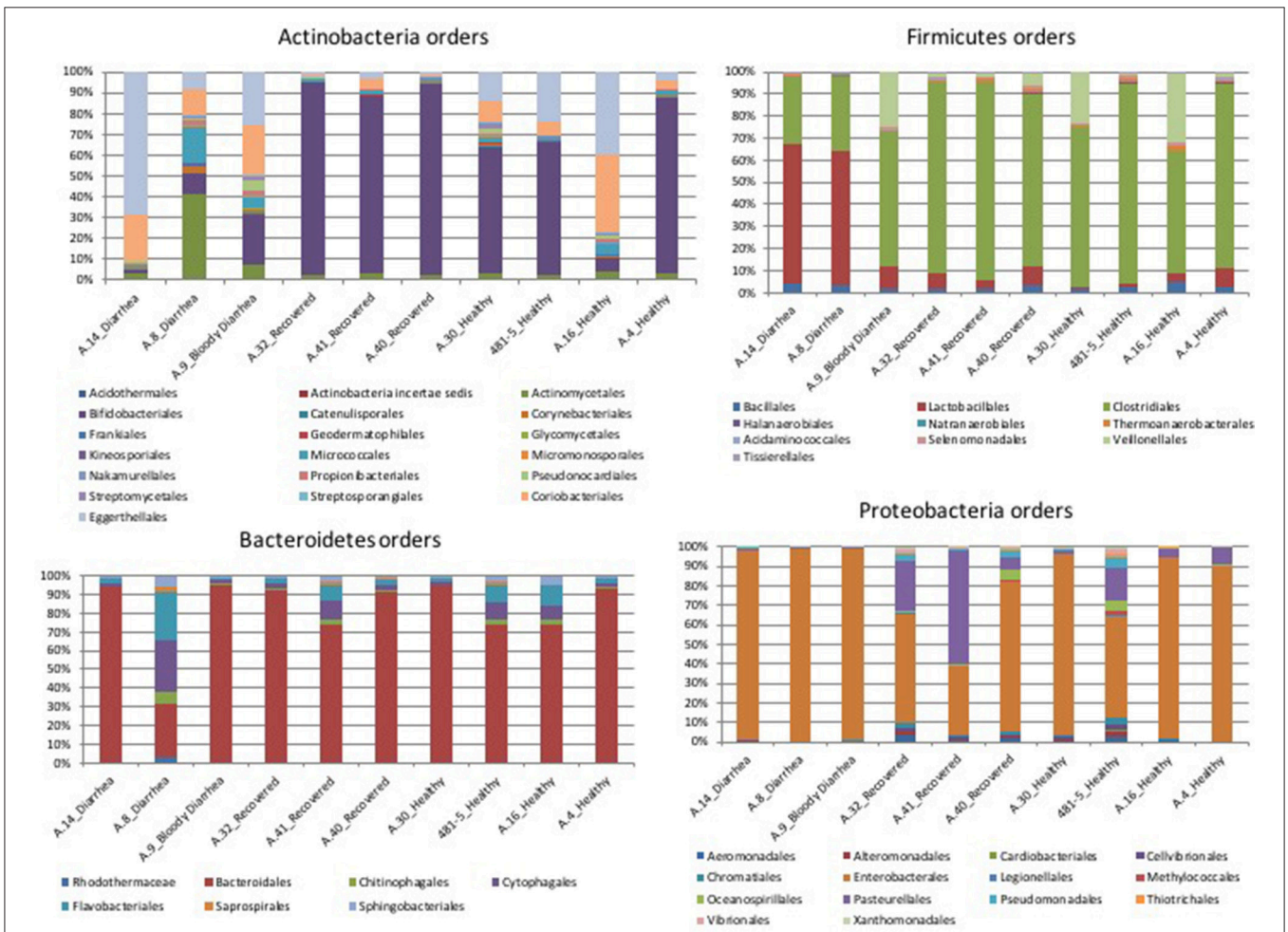


FIGURE 1 | Distribution of the OTUs in the metagenomes from the STEC O26 outbreak, obtained through the analysis based on read mapping. The scale on the y axis refers to the percentage of the reads mapping to the specific OTU.

involved in an outbreak caused by a STEC O26:H11 strain as well as from patients with Crohn's disease, with and without evidence of STEC and other pathogenic *E. coli* infections.

The identification of all the virulence features of the epidemic STEC strain in the metagenomes, previously determined through whole genome sequencing of the isolated outbreak strain (unpublished), showed that metagenomics is a promising diagnostic tool for infectious diseases, subjected to the availability of comprehensive databases of markers for the pathogens of interest. In sample A.14 the presence of the *iroN* gene, not present in the epidemic strain, was observed. The *iroN* gene encodes receptors involved in iron acquisition by bacteria (Russo et al., 2002), and it may have been present in a different co-circulating *E. coli* strain or even other bacteria.

The analyses carried out did not show the presence of genes encoding for Stx in specimens collected from both healed and

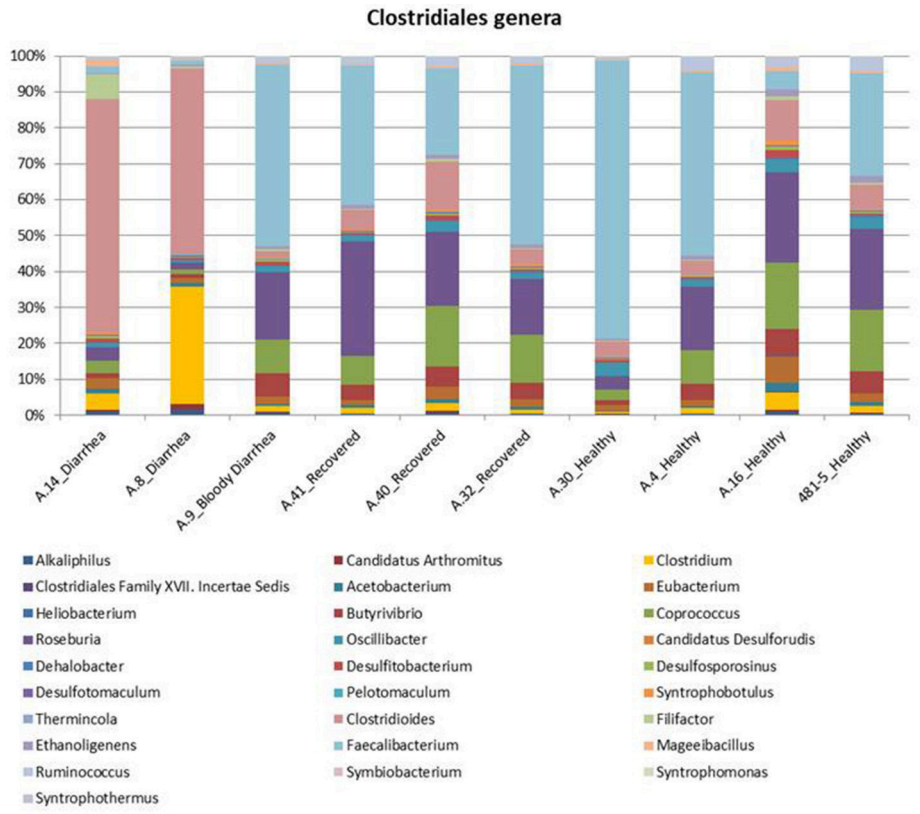


FIGURE 2 | Distribution of the main genera belonging to the *Clostridiales* order in the metagenomes from the STEC O26 outbreak, obtained through the analysis based on read mapping. The scale on the y axis refers to the percentage of the reads mapping to the specific OTU.

healthy subjects, indicating a good specificity of the approach used. Interestingly, in samples A.32, A.40, and A.41 (healed patients) evidence of a previous STEC infection was provided by the presence of some genes know to be located on the large virulence plasmid of STEC (**Table 1**). It is possible that this plasmid may have remained in the bacterial population after the STEC strain was cleared, leaving traces of the previous STEC infection.

The analysis of the metagenomes obtained from patients with Crohn's disease also confirmed previously obtained real-time PCR results (**Table 2**). For Sample 1 and Sample 5 this approach failed to identify the *escV* gene, associated with Enteropathogenic *E. coli* (EPEC). However, in these samples the evidence of EPEC was provided by the identification of the presence of other genes characteristic of this pathogroup, but different from those used in the real-time PCR experiments (**Table 2**). Indeed, while the real-time PCR approach may be more sensitive, due to the exponential amplification of the targets, the metagenomic approach may complement the lower technical sensitivity with the simultaneous search for more determinants associated with the same pathogroup. The use of a wider panel of targets may confer a stronger ability to detect pathogens, but it requires the availability of accurate and curated databases with the genetic markers characteristic for any of the pathogen of interest.

The taxonomic profiling of the STEC-infected samples provided insight into the changes occurring in the intestinal microbiota upon STEC infection. As previously described by other authors (Tap et al., 2009), our results showed that the *Bacteroidetes*, *Firmicutes*, *Proteobacteria*, *Actinobacteria* were the bacterial phyla that prevailed in the samples, confirming their proposed role in maintaining the homeostasis in the human intestine (Tap et al., 2009) (**Table 3**). A deeper taxonomic classification highlighted a marked prevalence of some members of the *Bifidobacteriales* and *Clostridiales* orders in the STEC negative samples (**Figures 1**, **2**, **4B**). It has been already described that members of the *Bifidobacterium* genus confer positive health benefits to the human host (O'Callaghan and van Sinderen, 2016). These bacteria exert a probiotic action, stimulating the immune system (Perdigon et al., 1995), and providing a protection against gastrointestinal pathogens colonization by competitive exclusion of enteropathogens based on common binding sites on epithelial cells (Gueimonde et al., 2007). In addition, Fukuda and colleagues demonstrated that acetate production by *B. longum* strains is linked to the *in vitro* protection of host epithelial cells from the effect of Shiga toxin (Fukuda et al., 2011). The high abundance of *B. longum* and *B. breve* species, and *Bifidobacterium* genus observed in the samples collected from healthy and healed patients (**Figure 4B**)

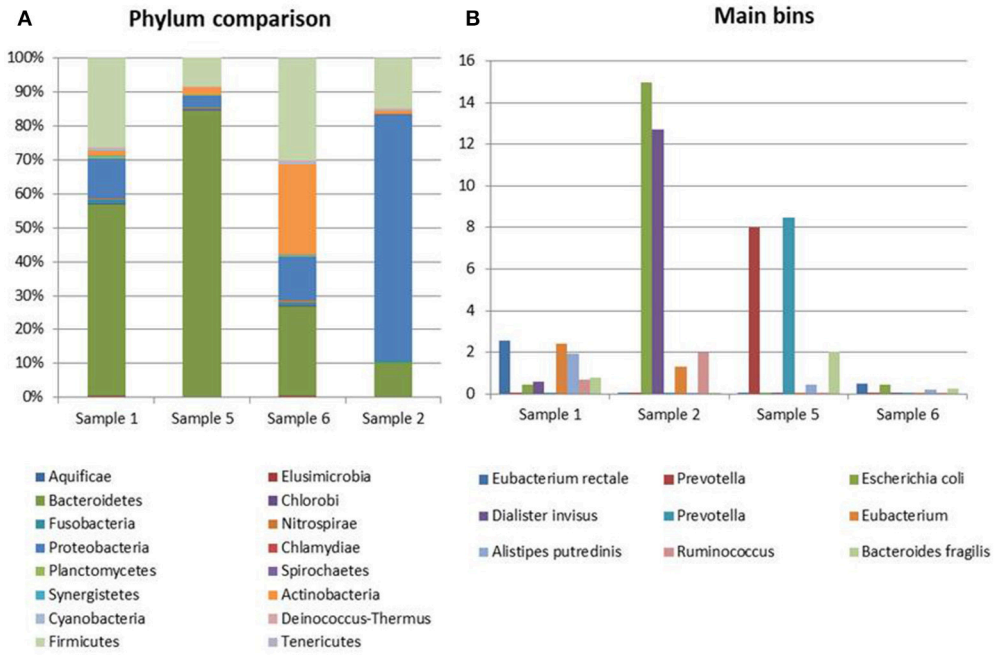


FIGURE 3 | Taxonomic profiling of the samples from CD patients. **(A)** Distribution of the main bacterial phyla in the metagenomes obtained through the analysis based on read mapping. **(B)** Distribution of bins with an abundance value >2% ($N = 10$).

could have different explanations. Such a high proportion of *Bifidobacterium* genus in the intestine may have been effective in protecting the healthy subjects against the STEC infection or may have favored the positive outcome of the infection in the healed patients, which have not developed the HUS. On the other hand, the high abundance observed of the members of this genus could have been the effect of a cleared or absent infection.

Similarly, the high abundance of the *Clostridiales* order in the STEC negative samples and the opposite prevalence of *Lactobacilalles* order in the STEC positive samples could be put into relation with the STEC infection. As a matter of fact, it is possible that the intestinal colonization by STEC may contrast the normal permanence of *Clostridia* in the intestine, favoring the presence of *Bacilli*. It is interesting to note that *Faecalibacterium*, *Roseburia*, *Coprococcus,* and *Butyrivibrio* genera, highly prevalent in the metagenomes belonging to the control group, have all been proved to have a beneficial role in the human host intestine (Hugon et al., 2017). In this study, we have included some samples from patients with Crohn's disease. It has been described that patients with CD show a global decrease in the biodiversity of the fecal microbiota, essentially due to a markedly reduced diversity of *Firmicutes* (Manichanh et al., 2006; Sokol et al., 2008). As a matter of fact, a significantly reduced abundance of *Roseburia*, *Faecalibacterium,* and other genera belonging to the *Clostridiales* order in the intestine of CD patients has been reported (Kang et al., 2010; Morgan et al., 2012; Gevers et al., 2014). We considered the CD samples as showing taxonomic profiles related with a general status of dysbiosis not related with STEC infection and have evaluated the differences observed between the samples from the cases and the controls collected in the framework of the STEC outbreak in the light of the profiles observed in the dysbiotic CD specimens. Our results confirmed the low biodiversity of the gut microbiota in CD patients and the lower proportion of members of the *Firmicutes* phylum (**Figure 3** and Figure S1). Additionally, our findings highlighted the absence of beneficial *Bifidobacterium* species in the feces collected from CD subjects, similarly to what was observed in the samples from STEC infections (**Figures 1**, **4**). This latter observation, confirms that a decrease in the abundance of *Bifidobacterium* is associated with an infection with diarrheagenic agents or a dysbiosis status as it has been previously proposed (Gevers et al., 2014; O'Callaghan and van Sinderen, 2016; He et al., 2017), suggesting that the perturbation in the proportions of the *Bifidobacterium* genus observed in this study may not be specific for the infection with STEC.

Our results indicate that metagenomics is effective in detecting genomic traits associated with STEC in stool samples from infected subjects, making it a promising tool for the culture-independent diagnosis of the infections. Additionally, the bioinformatic procedures used can be automated and applied to the detection of other pathogens. The analysis of the taxonomic composition of the intestinal microbiota showed a good agreement between the data obtained with both the reference-free and read mapping approaches, supporting the following comparative analyses. In this respect, to the best of our knowledge, our data provide a first evidence of the changes occurring in the intestinal microbiota of children in the course of

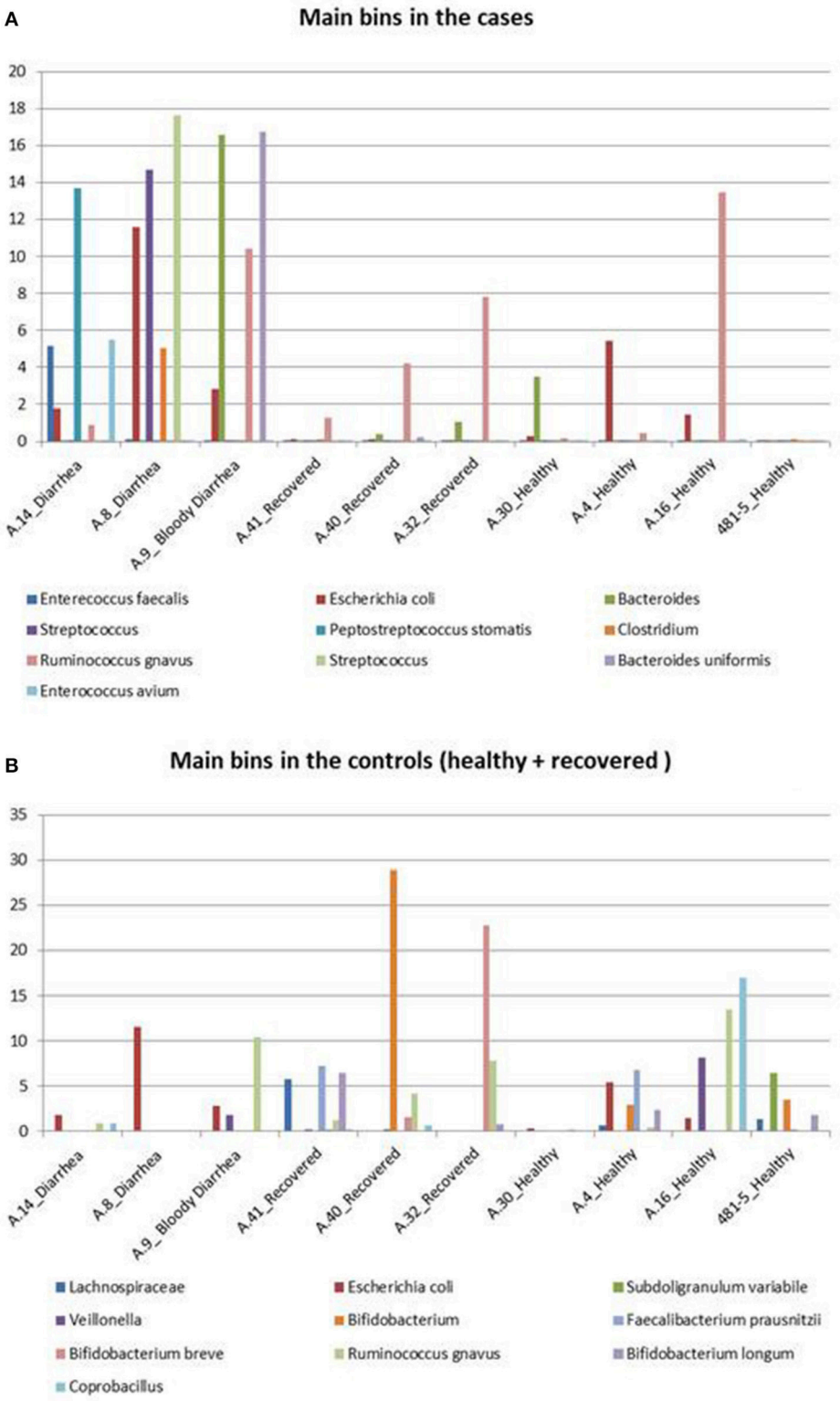


FIGURE 4 | Taxonomic profiling of the metagenomes from the STEC O26 outbreak using a reference-free approach. The figure highlights the bins identified in the cases and controls. **(A)** Shows the distribution of the 10 bins more abundant in the cases group across all the samples, while **(B)** illustrates the 10 bins which are more represented in the controls group, including healed and healthy subjects across all the samples.

STEC infection. Further studies are required to assess the reasons underlying such differences and if certain taxonomic profiles may be considered effective in protecting the host from acquiring the STEC infection.

AUTHOR CONTRIBUTIONS

FG performed the DNA extraction, the metagenomic sequencing, the bioinformatic analyses, and drafted the manuscript; FvM contributed to the bioinformatic analyses and revised the manuscript; AK installed the server for the bioinformatic analyses and provided assistance with the data analysis; VM contributed to the sequencing and bioinformatic analyses on the epidemic strain; GS contributed to the samples' seletion and critically revised the manuscript; FM contributed to the isolation of the epidemic strain; BD contributed to the bioinformatic analyses and revised the manuscript; HA collected the samples from Crohn' disease patients; GR contributed to the metagenomic sequencing; AF contributed to the metagenomic sequencing and revised the manuscript; JR contributed to the metagenomic sequencing and revised the manuscript; SM conceived the study and strongly contributed to revise the manuscript.

ACKNOWLEDGMENTS

BD was supported by the Netherlands Organization for Scientific Research (NWO) Vidi grant 864.14.004.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcimb.2018.00025/full#supplementary-material

REFERENCES

Armstrong, G. L., Hollingsworth, J., and Morris, 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. doi: 10.1093/oxfordjournals.epirev.a017914

Bäckhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A., and Gordon, J. I. (2005). Host-bacterial mutualism in the human intestine. *Science* 307, 1915–1920. doi: 10.1126/science.1104816

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Beitel, C. W., Froenicke, L., Lang, J. M., Korf, I. F., Michelmore, R. W., Eisen, J. A., et al. (2014). Strain- and plasmid-level deconvolution of a synthetic metagenome by sequencing proximity ligation products. *PeerJ* 2:e415. doi: 10.7717/peerj.415

Beutin, L., Bode, L., Ozel, M., and Stephan, R. (1990). Enterohemolysin production is associated with a temperate bacteriophage in *Escherichia coli* serogroup O26 strains. *J. Bacteriol.* 172, 6469–6475. doi: 10.1128/jb.172.11.6469-6475.1990

Brunder, W., Schmidt, H., and Karch, H. (1997). EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. *Mol. Microbiol.* 24, 767–778. doi: 10.1046/j.1365-2958.1997.3871751.x

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421

Cambuy, D. D., Coutinho, F. H., and Dutilh, B. E. (2016). Contig annotation tool CAT robustly classifies assembled metagenomic contigs and long sequences. *bioRxiv:072868*. doi: 10.1101/072868

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. doi: 10.1051/vetres:2005002

de Vos, W. M., and de Vos, E. A. (2012). Role of the intestinal microbiome in health and disease: from correlation to causation. *Nutr. Rev.* 70(Suppl. 1), S45–S56. doi: 10.1111/j.1753-4887.2012.00505.x

Forbes, J. D., Van Domselaar, G., and Bernstein, C. N. (2016). The gut microbiota in immune-mediated inflammatory diseases. *Front. Microbiol.* 7:1081. doi: 10.3389/fmicb.2016.01081

Fujiwara, S., Hashiba, H., Hirota, T., and Forstner, J. F. (2001). Inhibition of the binding of enterotoxigenic *Escherichia coli* Pb176 to human intestinal epithelial cell line HCT-8 by an extracellular protein fraction containing BIF of *Bifidobacterium longum* SBT2928: suggestive evidence of blocking of the binding receptor gangliotetraosylceramide on the cell surface. *Int. J. Food Microbiol.* 67, 97–106. doi: 10.1016/S0168-1605(01)00432-9

Fukuda, S., Toh, H., Hase, K., Oshima, K., Nakanishi, Y., Yoshimura, K., et al. (2011). Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* 469, 543–547. doi: 10.1038/nature09646

Gamage, S. D., Strasser, J. E., Chalk, C. L., and Weiss, A. A. (2003). Nonpathogenic *Escherichia coli* can contribute to the production of Shiga toxin. *Infect. Immun.* 71, 3107–3115. doi: 10.1128/IAI.71.6.3107-3115.2003

Garza, D. R., and Dutilh, B. E. (2015). From cultured to uncultured genome sequences: metagenomics and modeling microbial ecosystems. *Cell. Mol. Life Sci.* 72, 4287–4308. doi: 10.1007/s00018-015-2004-1

Gauthier, A., Puente, J. L., and Finlay, B. B. (2003). Secretin of the enteropathogenic *Escherichia coli* type III secretion system requires components of the type III apparatus for assembly and localization. *Infect. Immun.* 71, 3310–3319. doi: 10.1128/IAI.71.6.3310-3319.2003

Gevers, D., Kugathasan, S., Denson, L. A., Vázquez-Baeza, Y., Van Treuren, W., Ren, B., et al. (2014). The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* 15, 382–392. doi: 10.1016/j.chom.2014.02.005

Gotelli, N. J., and Colwell, R. K. (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.* 4, 379–391. doi: 10.1046/j.1461-0248.2001.00230.x

Gueimonde, M., Margolles, A., de los Reyes-Gavilán, C. G., and Salminen, S. (2007). Competitive exclusion of enteropathogens from human intestinal mucus by Bifidobacterium strains with acquired resistance to bile–a preliminary study. *Int. J. Food Microbiol.* 113, 228–232. doi: 10.1016/j.ijfoodmicro.2006.05.017

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086

He, Q., Gao, Y., Jie, Z., Yu, X., Laursen, J. M., Xiao, L., et al. (2017). Two distinct metacommunities characterize the gut microbiota in Crohn's disease patients. *Gigascience* 6, 1–11. doi: 10.1093/gigascience/gix050

Hugon, P., Lagier, J. C., Colson, P., Bittar, F., and Raoult, D. (2017). Repertoire of human gut microbes. *Microb. Pathog.* 106, 103–112. doi: 10.1016/j.micpath.2016.06.020

Huson, D. H., Auch, A. F., Qi, J., and Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Res.* 17, 377–386. doi: 10.1101/gr.5969107

Joensen, K. G., Scheutz, F., Lund, O., Hasman, H., Kaas, R. S., Nielsen, E. M., et al. (2014). Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *J. Clin. Microbiol.* 52, 1501–1510. doi: 10.1128/JCM.03617-13

Kang, D. D., Froula, J., Egan, R., and Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 3:e1165. doi: 10.7717/peerj.1165

Kang, S., Denman, S. E., Morrison, M., Yu, Z., Dore, J., Leclerc, M., et al. (2010). Dysbiosis of fecal microbiota in Crohn's disease patients as revealed by a custom phylogenetic microarray. *Inflamm. Bowel Dis.* 16, 2034–2042. doi: 10.1002/ibd.21319

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Ley, R. E., Peterson, D. A., and Gordon, J. I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124, 837–848. doi: 10.1016/j.cell.2006.02.017

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595. doi: 10.1093/bioinformatics/btp698

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., et al. (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* 55, 205–211. doi: 10.1136/gut.2005.073817

Morgan, X. C., Tickle, T. L., Sokol, H., Gevers, D., Devaney, K. L., Ward, D. V., et al. (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol.* 13:R79. doi: 10.1186/gb-2012-13-9-r79

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 27, 824–834. doi: 10.1101/gr.213959.116

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

O'Callaghan, A., and van Sinderen, D. (2016). Bifidobacteria and their role as members of the human gut microbiota. *Front. Microbiol.* 7:925. doi: 10.3389/fmicb.2016.00925

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

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114

Perdigon, G., Alvarez, S., Rachid, M., Agüero, G., and Gobbato, N. (1995). Immune system stimulation by probiotics. *J. Dairy Sci.* 78, 1597–1606. doi: 10.3168/jds.S0022-0302(95)76784-4

Rivière, A., Selak, M., Lantin, D., Leroy, F., and De Vuyst, L. (2016). Bifidobacteria and butyrate-producing colon bacteria: importance and strategies for their stimulation in the human gut. *Front. Microbiol.* 7:979. doi: 10.3389/fmicb.2016.00979

Russo, T. A., McFadden, C. D., Carlino-MacDonald, U. B., Beanan, J. M., Barnard, T. J., and Johnson, J. R. (2002). IroN functions as a siderophore receptor and is a urovirulence factor in an extraintestinal pathogenic isolate of *Escherichia coli*. *Infect. Immun.* 70, 7156–7160. doi: 10.1128/IAI.70.12.7156-7160.2002

Shreiner, A. B., Kao, J. Y., and Young, V. B. (2015). The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.* 31, 69–75. doi: 10.1097/MOG.0000000000000139

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermúdez-Humarán, L. G., Gratadoux, J. J., et al. (2008). *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl. Acad. Sci. U.S.A.* 105, 16731–16736. doi: 10.1073/pnas.0804812105

Tap, J., Mondot, S., Levenez, F., Pelletier, E., Caron, C., Furet, J. P., et al. (2009). Towards the human intestinal microbiota phylogenetic core. *Environ. Microbiol.* 11, 2574–2584. doi: 10.1111/j.1462-2920.2009.01982.x

Tatsuno, I., Kimura, H., Okutani, A., Kanamaru, K., Abe, H., Nagai, S., et al. (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. doi: 10.1128/IAI.68.10.5943-5952.2000

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.

Supplementary Material

Table S1. Assembly statistics indicating the quality of the cross-assembled scaffolds evaluated using QUAST.

Statistics without reference	CrossIllumina scaffolds	CrossIllumina scaffolds broken
# contigs	429,862	432,997
# contigs (>= 0 bp)	1,461,986	-
# contigs (>= 1000 bp)	137,737	140,630
# contigs (>= 5000 bp)	24,314	24,636
# contigs (>= 10000 bp)	11,152	10,968
# contigs (>= 25000 bp)	3,316	3,123
# contigs (>= 50000 bp)	1,050	982
Largest contig	505,337	505,337
Total length	797,120,371	795,872,475
Total length (>= 0 bp)	1,127,506,476	-
Total length (>= 1000 bp)	605,109,732	603,691,131
Total length (>= 5000 bp)	384,370,951	376,534,562
Total length (>= 10000 bp)	293,205,960	281,978,519
Total length (>= 25000 bp)	174,419,520	163,960,235
Total length (>= 50000 bp)	96,898,633	91,095,529
N50	4,483	4,265
N75	1,043	1,040
L50	27,313	29,272
L75	1,30,617	1,33,973
GC (%)	44.87	44.87
Mismatches		
# N's	11,29,634	128
# N's per 100 kbp	141.71	0.02

Table S2. Number and percentage of the reads from the 14 samples analysed in this study mapping against the cross-assembly. Samples A.9, A.4 and A.30 were sequenced with the Ion Torrent platform.

	Number of the mapping reads	Percentage of the mapping reads
A.9	3856842	90.93%
A.8	11936333	98.95%
A.14	12250589	98.82%
A.32	14622057	99.59%
A.40	11348934	99.60%
A.41	13067306	99.35%
481-5	8002827	99.20%
A.4	2931083	87.13%
A.30	3382512	85.96%
A.16	16932979	99.71%
Sample 1	20107751	99.63%
Sample 2	15901940	99.46%
Sample 5	19003584	99.77%
Sample 6	12079692	99.23%

Figure S1. Rarefaction plot of the 14 metagenomes analysed in this study. The curves on the top of the chart correspond to the samples from the STEC O26 outbreak; while those present on the bottom part correspond to the samples from CD patients.

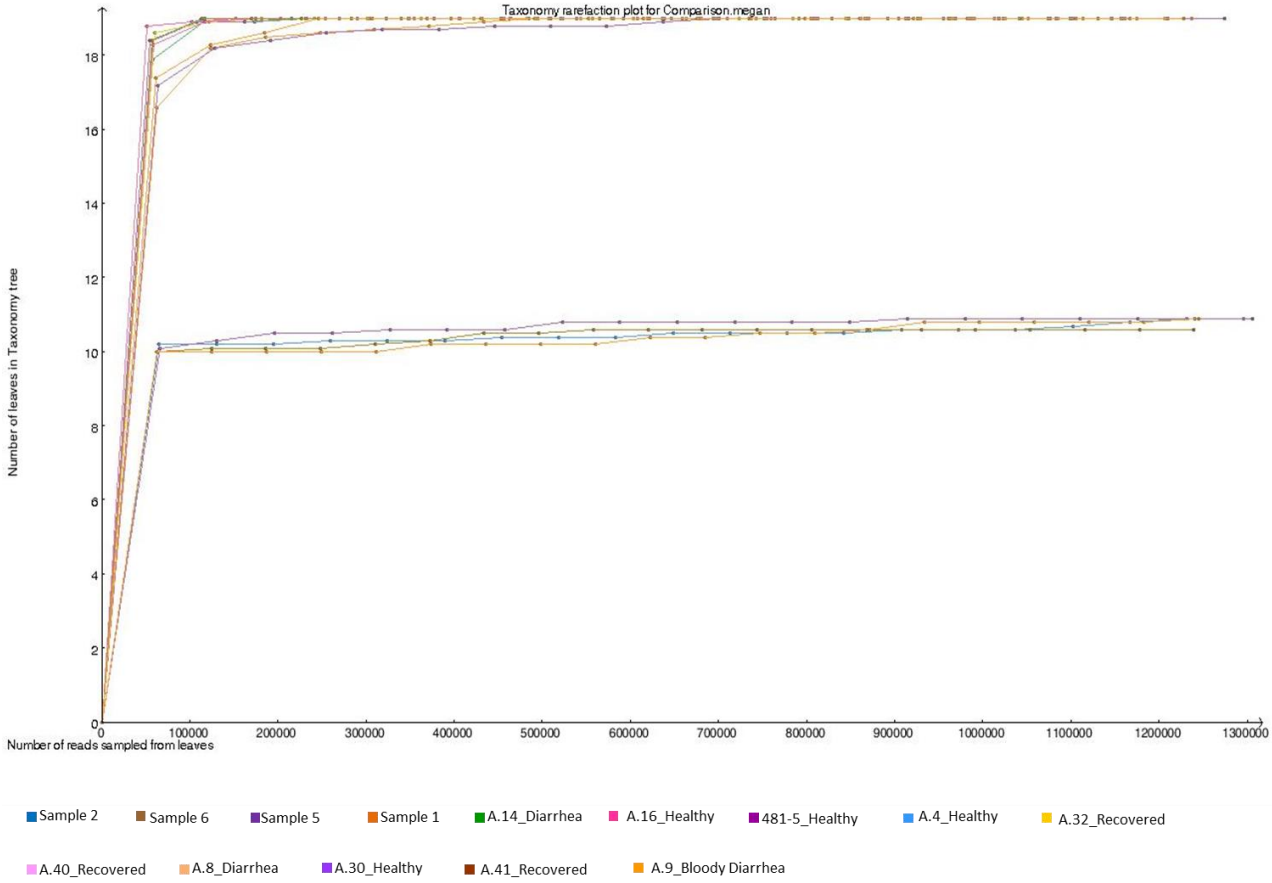
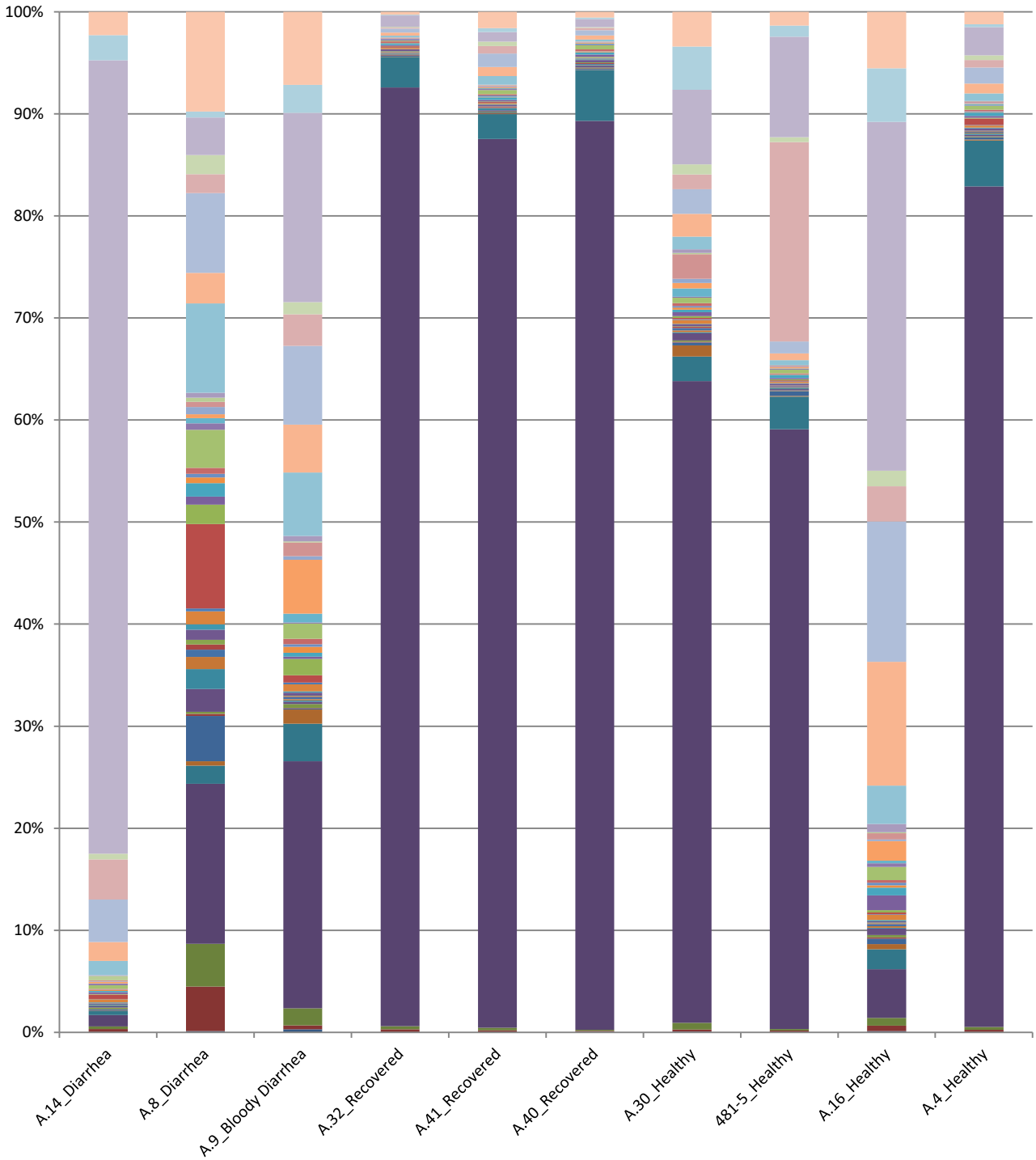


Figure S2. Distribution of the species belonging to *Actinobacteria* phylum in the metagenomes from the STEC O26 outbreak, obtained through the analysis based on read mapping. The scale on the y axis refers to the percentage of the reads mapping to the specific OTU.

Actinobacteria species



- Thermobispora bispora DSM 43833
- **Bifidobacterium longum**
- Corynebacterium glutamicum
- Beutenbergia cavernae
- Jonesia denitrificans
- Microbacterium testaceum
- Renibacterium salmoninarum
- Xylanimonas cellulosilytica
- Verrucosispora maris
- Microlunatus phosphovorus
- Kitasatospora setae
- Streptosporangium roseum
- Coriobacterium glomerans
- Eggerthella lenta
- Arcanobacterium haemolyticum
- Gardnerella vaginalis
- Nocardia brasiliensis
- Cellulomonas fimi
- Clavibacter michiganensis
- Kocuria rhizophila
- Rothia dentocariosa
- Sanguibacter keddieii
- Nakamurella multipartita
- Amycolatopsis mediterranei
- Streptomyces bingchenggensis
- Atopobium parvulum
- Adlercreutzia equolifaciens
- Gordonibacter pamelaeae
- Mobiluncus curtisii
- Catenulispora acidiphila
- Modestobacter marinus
- Brachybacterium faecium
- Leifsonia xyli
- Pseudarthrobacter chlorophenolicus
- Isoptericola variabilis
- Actinoplanes friuliensis
- Acidipropionibacterium acidipropionici
- Saccharopolyspora erythraea
- Nocardiopsis alba
- Olsenella uli
- Cryptobacterium curtum
- Slackia heliotrinireducens

Chapter 7

Discussion and Concluding remarks

7. Discussion and Concluding remarks

Shiga toxin-producing *E. coli* are a heterogeneous group of zoonotic pathogens causing human diseases with a wide spectrum of clinical manifestations. The major reservoir of STEC are ruminants, especially the cattle, which are asymptomatic and represent one of the most important source of human infections. The principal transmission route of STEC is the ingestion of contaminated food of animal origin. Faecal contamination of foodstuffs of animal origin can occur frequently during several steps of food production and undercooked meat, or milk and dairy products not subjected to an efficient pasteurization process, represent the food commodities mainly involved in human infections. Other transmission pathways include the inter-human cycle and direct contacts with animals carrying STEC strains (Henderson, 2008, Ihekweazu *et al.*, 2012) as well as other vehicles, including recreational and drinking water and food of vegetable origin (Caprioli *et al.*, 2005). Over the years, the concern related to foodborne outbreaks associated with the consumption of fresh produce increased. The soil used for growing crops is often treated with manure and irrigated with water that can transport pathogenic microorganisms, resulting in surface contamination of vegetables and fruits.

The dissemination of STEC and other pathogenic *E. coli* in farmland is augmented by the use in agriculture of organic matter produced at wastewater treatment plants from municipal and zootechnical sewage sludge. Such sludges usually are not treated by disinfection processes to remove biological pollutants (Mezzanotte *et al.*, 2007), allowing the delivery of pathogenic *E. coli* strains to the soils where they can persist for long times, representing a potential risk of microbial contamination of the food chain.

The farm environment can also be considered a reservoir of STEC, where they can survive up to many months (Gagliardi *et al.*, 2002, Fremaux *et al.*, 2008), causing a continuous exposure of grazing and wild animals and triggering amplification cycles.

Once they reach the human host, STEC are capable to cause infections with a wide spectrum of clinical signs including very severe forms, as the haemorrhagic colitis (HC) and the haemolytic uremic syndrome (HUS). STEC pathogenicity mainly depends on the presence of *stx*-converting bacteriophages, that confer the capability to produce the Shiga toxins. However, the presence of *stx*-phages seems to be not sufficient to cause HC or HUS and additional virulence factors are required (Karch, 2001). Such factors are mainly related with an efficient colonization of the host intestinal mucosa. One of the best known of such mechanisms is the attaching and effacing (A/E) lesion. The A/E mechanism is governed by genes located in the LEE pathogenicity island (Jerse *et al.*, 1990). Although the presence of LEE locus is associated to the development of HUS, STEC from diverse

serogroups that do no encode the LEE PAI have been frequently isolated from patients with severe clinical complications (Johnson *et al.*, 2006, Newton *et al.*, 2009). These strains use several other virulence genes carried by mobile genetic elements (MGE).

As a matter of fact, the transfer of MGE between different *E. coli* strains or even derived from other bacterial species represents the basis of the great diversity of STEC strains. These MGE carry accessory genomic information and constitute the 26% of the total *E. coli* pangenome (Touchon *et al.*, 2009, van Elsas *et al.*, 2011) (Figure 7.1) and it has been described that MGE account for about 20% of the whole genome of STEC O157:H7 (Perna *et al.*, 2001).

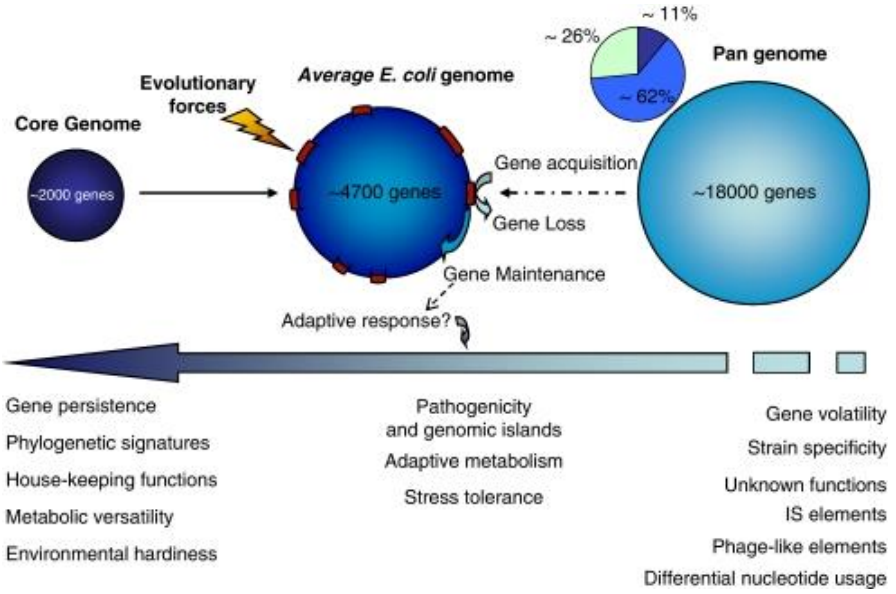


Figure 7.1 Representation of *E. coli* pangenome constituted of about 18,000 genes, of which 11% form the conserved part, the "core genome" (dark blue), 62% represents the "persistent genes" (blue), 26% can be considered "accessory genes" (pale blue). Evolutionary forces induce acquisition and loss of genetic determinants, leading to an average of *E. coli* genome of about 4,700 genes (van Elsas *et al.*, 2011).

As far as the pathogenicity of STEC is concerned, it has to been considered that the virulence genes content of the strain causing the infection is necessary but may not be sufficient, as other factors may have important roles in the progression of the symptoms (Chapter 6).

7.1 Stx2f producing-*E. coli* represent a serious problem in public health

Shiga toxins (Stxs) are the main virulence determinants associated with STEC. The genes encoding the Stxs are harboured by temperate lambdoid bacteriophages (O'Brien *et al.*, 1984). Differences in

the DNA sequence of the *stx* genes led to classify the Stxs in two main antigenically diverse types, Stx1 and Stx2, which in turn are subdivided into three (Stx1a, Stx1c and Stx1d) and seven subtypes (Stx2a-g) respectively (Persson *et al.*, 2007).

The most severe diseases in human have been frequently associated to the toxin subtype produced by the infecting strains (Friedrich *et al.*, 2002). For long time STEC producing Stx2f subtype have been considered of scarce epidemiological relevance, as they were considered adapted to pigeons, their natural reservoir, and rarely associated with human infections. Therefore, cases of infections by Stx2f-producing *E. coli* have been long ignored and not efficiently surveyed.

Originally, only a few sporadic cases of mild diarrhoea associated to the Stx2f variant were reported in literature, but nowadays an increasing number of *stx2f* positive strains, isolated from human disease and belonging to a great variety of different serotypes, have been described (Prager *et al.*, 2009, Friesema *et al.*, 2014), suggesting that a more thorough assessment of the risk associated with this STEC population was needed.

We have carried out a whole genome characterization of *E. coli* strains producing the Stx2f subtype including isolates from diarrhoea, HUS and the natural reservoir, in order to provide evidence on the possible origin of the Stx2f-producing STEC causing HUS. Virulotyping of the isolates highlighted the existence of distinct subpopulations of these strains, which suggested a general variability of such STEC population was not identified previously. In particular, the isolates grouped based on their virulence genes content and with different degrees of overlapping with those present in the animal reservoir. The same analysis showed that the isolates from HUS cases displayed an asset of virulence genes additional to the *stx* resembling that of the more typical STEC producing Stx2 subtypes, such as the Stx2a, normally associated with HUS and belonging to serogroups such as O157, O26, O111, O103 and O145. Beside the chromosomally encoded genes governing the A/E lesion, and the related accessory factors such as the non-LEE encoded effectors (nle), encoded by the *nleABC* genes (Garmendia *et al.*, 2005), the Stx2f-producing STEC from HUS also carried the OI-122 pathogenicity island (Karmali *et al.*, 2003), and the genes *ehxA*, *espP* and *katP* genes, present on the large virulence plasmid first described in STEC O157 (pO157) (Beutin *et al.*, 1990, Brunder *et al.*, 1996, Brunder *et al.*, 1997). Our results allowed us to speculate on the potential association of the Stx subtypes with the severity of the symptoms. As a matter of fact, our study indicated that even though the Stx2f was previously considered as being adapted to the pigeon host and not involved in the induction of severe disease in humans, it still has the potential to cause HUS when produced by an isolate able to efficiently colonize the human intestinal tract.

In this respect, it is important to note that the Stx2f-STEC from pigeons and mild diarrhoea specimens were all positive for the presence of the LEE locus, but the latter lacked of all the

accessory determinants such as the *nle* and the genes conveyed by the OI-122 and pO157. At the same time the pigeon isolates showed an intermediate virulence genes asset with the presence of *nle* but did not possess the other two mentioned MGEs.

The compartmentalisation of the different subpopulations of Stx2f-producing STEC was also observed through the principal component analysis of virulence genes profile and a phylogenetic analysis conducted considering the whole genome SNPs showed that HUS Stx2f-producing strains clustered closer to other non-O157 STEC isolated from humans with severe disease, while the isolates from diarrhoea were more distantly related.

Whether the Stx2f-producing STEC isolated from HUS originated directly in the animal reservoir or through a mechanism of phage acquisition occurred in the environment and involving *E. coli* strains with a complete machinery for the induction of the A/E lesion, remains to be ascertained.

Further studies are also required to investigate the mechanism used by phages conveying *stx2f* genes to be released from the bacterial cell. In fact, these phages seem to lack most of the genes that regulate the switch between the lysogenic state and the lytic cycle typical of the lambdoid bacteriophages, as *cro*, *cI*, *cII* and *cIII* and we were not able to induce them in different experimental conditions (data not shown). Moreover, it will be interesting to assess the presence of free *stx2f*-paghes in the environment, in order to investigate whether these can be available to lysogenize other *E. coli* strain of non-pigeon sources.

7.2 "Bio-waste" as a new possible transmission pathway of pathogenic *E. coli* including STEC

STEC infections occurs, mainly, via foodborne and waterborne routes, through inter-human cycles or direct contacts with STEC positive animals, or their environment. Contamination of water and soil used for agricultural use by STEC and other pathogenic *E. coli* is gaining increasing attention in public health.

Nowadays, the need to high crop with a reduced usage of water and expensive fertilizers to amend soils led to recycle bio-waste as a resource for energy, water and nutrients (Figure 7.2) (Nielsen, 2017). Such waste, defined also as Biosolid (BSO), refers to sewage sludge derived from biological or chemical treatment of industrial, municipal and zootechnical wastewater, manure from livestock, and compost from green wastes (Saveyn *et al.*, 2014). Sewage sludge is subjected to both aerobic and anaerobic fermentation processes reaching high temperature, which, however, may not enough to remove potential hazards for human and animal health (Figure 7.2).

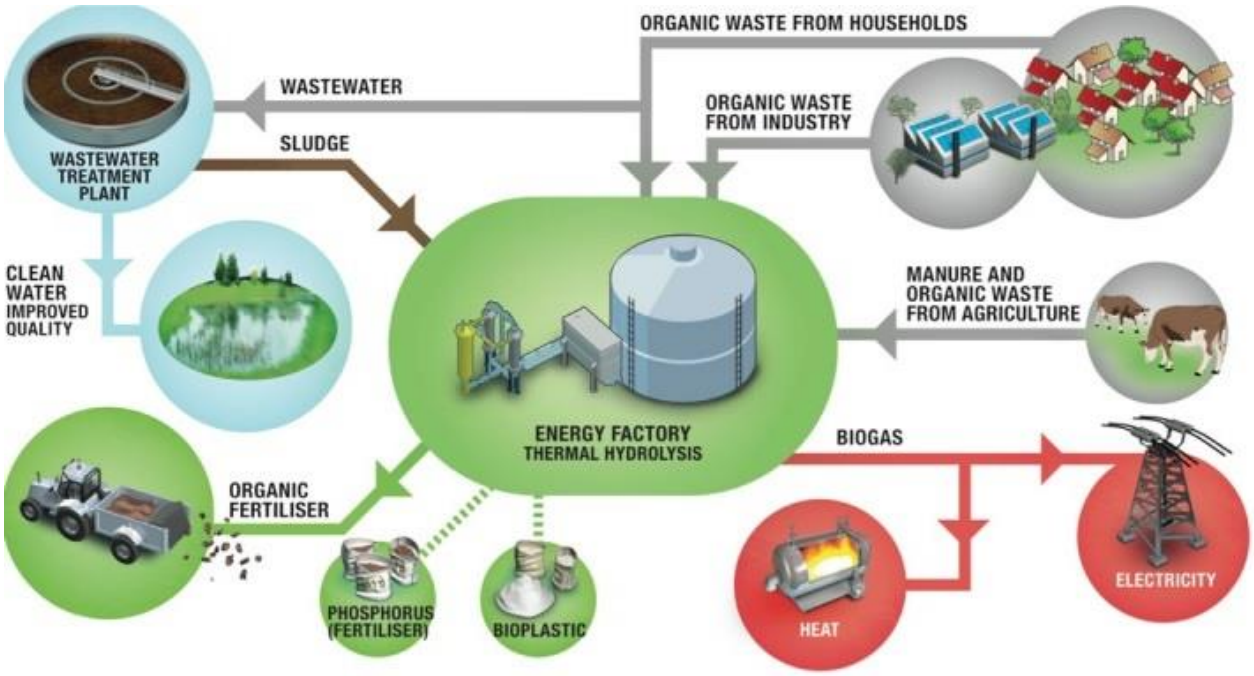


Figure 7.2 Schematic representation of recovery cycle and reuse of wastewater http://www.billundbiorefinery.dk/en/.

We attempted to identify free *stx2f*-phages (See 7.1) in BSO samples used for the analysis described in Chapters 4 and 5 (data unpublished). To this purpose we selected BSO samples positive for the presence of *stx2f* gene subtype in Real Time PCR to infect a bacterial culture with the filtered supernatant. The results showed the presence of many lytic plaques (Figure 7.3), indicating the presence of infective phages. We tried to identify possible *stx2f*-phages among the plaques through plaque hybridisation with a probe specific for the *stx2f* gene, but with negative results.

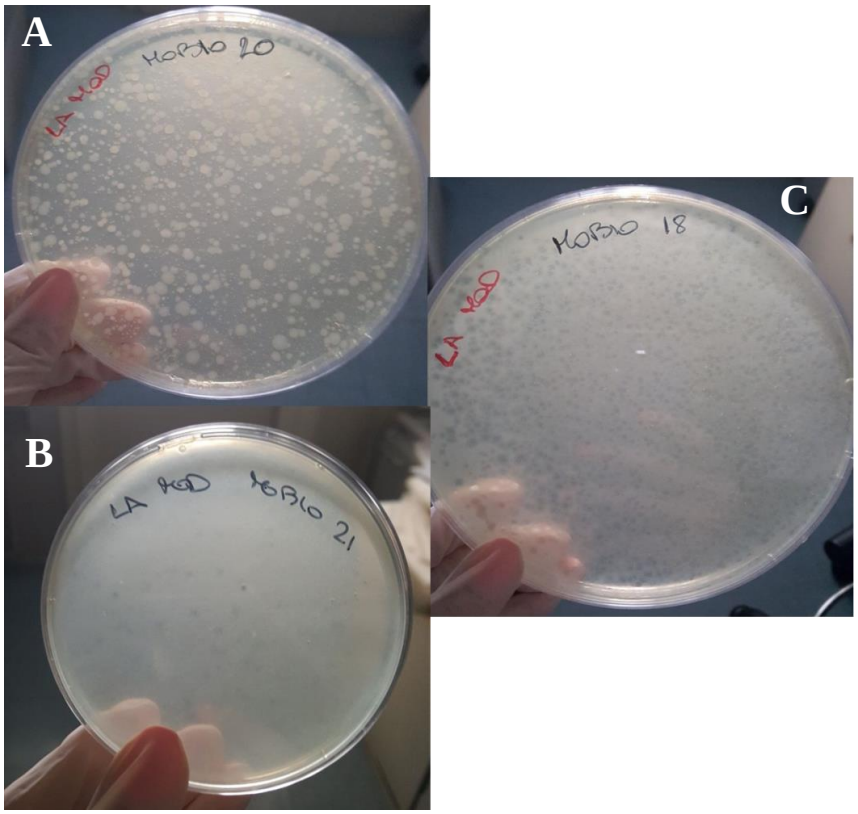


Figure 7.3 Agar plates containing phages plaques originated by lysis of the *E. coli* propagator strain upon infection with filtered supernatant derived from BSO samples.

We have also explored the possibility that BSO may be a source of pathogenic *E. coli* including STEC. We detected the presence of virulence genes associated with both viral and bacterial enteric pathogens in various BSO samples intended for soil fertilization and grazing activities. The analysis, performed using Real Time PCR, gave positive results for virulence genes of human adenovirus, human norovirus and different diarrhoeagenic *E. coli* (DEC). The screening for pathogenic *E. coli*-associated genes was performed on the DNA extracted both directly from the specimens and from enriched cultures of the samples. As described in Chapter 4, some Real Time PCR signals, which were not identified in the DNA extracted directly from the samples, appeared after the enrichment for some of the virulence genes investigated, suggesting the presence of viable microbial cells. Five specimens negative for *stx* genes in direct screening, became positive after the enrichment, indicating that biosolids may indeed be a potential source of STEC and their use in agriculture may cause the transmission of such important human pathogens to soils, with a consequent risk of contamination of the food chain. As a matter of fact, it has been demonstrated

that viable *E. coli* O157:H7 strains can persist up to several months in the soil (Gagliardi *et al.*, 2002).

Overall, the BSO samples analysed proved positive for the presence of virulence genes associated to all the *E. coli* pathotypes forming the DEC group. As a consequence, the genes investigated, all borne on mobile genetic elements (MGE), may be transferred to soils also in the absence of live bacterial cells and may be acquired by soil microorganisms, or by commensal species, including *E. coli*, living in the intestine of grazing animals, contributing to the possible emergence of new pathogenic variants.

One of the main concerns regards the emergence of novel STEC strains with shuffled virulence features. Bacteriophages harbouring *stx* genes can be acquired by *E. coli* strains belonging to different pathogroups, favouring the possible emergence of pathogenic clones highly virulent to human (Frank *et al.*, 2011, Tozzoli *et al.*, 2014). As a matter of fact, the samples analysed were from different origin including wastewater plants treating both wastes from urban settlement and from livestock and, although we could not isolate the strain responsible for the PCR signals, probably due to the massive amount of microorganisms present in the specimens, the possibility exists that such practice of using BSO from wastewater treatment plants to fertilize agricultural soils may represent a risk for human health and should be better investigated.

In order to come to a deeper analysis of the BSO samples we used metagenomics as described in Chapter 5. In particular, we used this approach to define a complete microbiological profile of BSO specimens including samples from sludges and compost. Interestingly, the taxonomic profiling of the BSO microbiota has proven to be a good indicator of the origin of the different BSO, with samples from sludges grouping together and apart from those from compost.

The metagenomic analysis provided evidences that biosolids may act as a new environmental vehicle of important biological threats for human health. The presence of many genomics traits related with different diarrheagenic *E. coli* pathogroups was detected, including STEC, confirming previous results obtained for the same samples by Real Time PCR (Chapter 4). The holistic analysis conducted in this study through the shotgun metagenomics also allowed us to identify the presence in the BSO samples of gene sequences associated with the resistance to the main classes of antibiotics, as Fluoroquinolones, Beta-lactams, Streptothricin, Fosfomycin, Vancomycin and Methicillin, as well as with the resistance to the heavy metals. This finding is in line with the derivation of the BSO from sludges from wastewater treatment plants. As a matter of fact, antimicrobials are used both in the treatment of human diseases and in animal farming and thus our results may reflect the pressure applied to bacteria from human and animal origin.

The metagenomic functional profiling of BSO may also provide an interesting tool for both retrospective and prospective analysis of emerging hazards. This may allow sound time-trends analysis to assess potential risks for animal and human health, as in the case of a screening for carbapenemase genes. Such genes were described to confer resistance to Carbapenems in environmental *Enterobacteriaceae* (Nordmann *et al.*, 2011, Tzouvelekis *et al.*, 2012). As a matter of fact, Carbapenems represent the last generation of Beta-lactams to fight multi-resistant bacteria responsible for serious infections at hospitals. However, there is rising evidence that wastewater effluents from hospital may play a role in disseminating Carbapenems resistance in the environment, through the genes-exchange between bacterial hosts in sewage (Ludden *et al.*, 2017). Although we did not carry out a specific analysis for this particular trait in the samples, it is important to note that one of the strength of the approach adopted is that the metagenomes are always available for further investigations and, given the huge amount of information contained into them, can be re-analysed with different bioinformatics procedures for specific aims.

7.3 Characterization of human intestinal microbiota in presence of STEC infections

The previous work on BSO metagenomic profiling demonstrated how the advent of next generation technologies changed the paradigm of the investigation of complex systems. In fact, it allows performing high-throughput analysis, enhancing the identification of a huge variety of different microorganisms and their associated genes. Based on these considerations, we decided to develop an improved metagenomic approach, based on the reconstruction from the metagenomes of putative entire genomes, in order to evaluate possible perturbations of the intestinal microbiota in human patients suffering from STEC infection. It has been proposed that alterations of intestinal microbial community, for instance due to the use of antibiotics, increased the susceptibility to gut colonization by enteric pathogens (Lagier, 2016). Several studies also proposed that commensal bacterial species play a crucial role in human health and may hinder the gastrointestinal tract colonization by pathogens, through different mechanisms (Rolhion *et al.*, 2016). This background information was the rationale to investigate whether changes in the intestinal microbiota following infection with a STEC strain could be recognised in order to better understand the dynamics of the pathogenesis and eventually open the way to mitigate the impact of STEC infections which cannot benefit from antibiotic treatment.

To this purpose we have selected samples for the metagenomic analysis collected during an outbreak of STEC O26 infections at a chid care center. This setting was considered as being the most appropriate as all the patients and controls were of the same age and all lived in the same area, reducing the bias of macroscopic differences due to different degrees of maturation of the

microbiota or to the provenience of subjects sampled. Our analyses allowed us to obtain a picture of the intestinal microflora variations in presence of STEC in comparison to healthy or healed subjects. Generally, the microbial distribution in specimens from diseased patients, showed a lesser abundance of commensal species belonging to the orders of *Bifidobacteriales* and *Clostridiales*. It is known that bacterial species belonging to these orders, such as *Bifidobacterium longum* and *Faecalibacterium prausnitzii*, are considered as providing a beneficial effect in contrasting intestinal infections (O'Callaghan *et al.*, 2016, Hugon *et al.*, 2017). As a matter of fact, in the complex scenario related to the interactions between the gut microbiota and the enteric pathogens, the abundance of commensal bacteria in the intestine of a healthy subject could hinder the success of STEC in achieving an efficient colonization of the host gut limiting the severity of symptoms. It has to be observed, however, that the decrease in the abundance of *Bifidobacterium* spp. may not be a specific variation associated with the infection with STEC. Perturbation in the proportion of *Bifidobacteriales* and *Crostridiales* orders has been already described to occur in presence of other diarrheagenic agents (O'Callaghan *et al.*, 2016) and also related with a general status of dysbiosis as observed in complex intestinal afflictions as the Crohn's disease (Chapter 6). In fact, it is known, that Crohn's disease patients display a general reduction in the complexity of the intestinal microbiota, favouring the colonization by pathogenic microorganisms (de Vos *et al.*, 2012).

The findings obtained in this pilot study represent the first evidence of how the intestinal microbiota reacts to the infection with a STEC O26:H11 and is a first attempt to understand the dynamics establishing in the host intestine upon infection. A more complete comprehension of microbiota-STEC interaction, e.g. by studying the changes occurring following infection with other STEC serotypes, would be of great help to unravel the host-related mechanisms of pathogenesis. Indeed, STEC strains presenting the same virulence genes background appear involved in a wide range of clinical symptoms in human host, ranging from mild diarrhoea to the life-threatening HUS as well as STEC with different genomic assets can be isolated from HUS cases. These investigations may ultimately lead to new approaches to mitigate the burden of STEC infections or to alleviate the symptoms induced upon infection.

7.4 Concluding remarks

The work described in this PhD thesis developed throughout the entire axis of exposure to STEC infections, from the animal reservoir to the development of disease. Given the complex epidemiology of this infectious disease, the proposed work has been focused on the less known segments of this axis, as the investigation of STEC isolates whose impact on public health was not completely clear as well as on the identification of new reservoirs and vehicle of dissemination of

the pathogens throughout the whole food chain, up to the less known action of the STEC on the host intestinal microbiota.

The next generation sequencing technologies, extensively used in the studies carried out, provided a wealth of information on all the topics touched and allowed to recognize new STEC types causing HUS, new possible pathways for the spreading of STEC and other pathogenic *E. coli* and the risk of their transfer to the food chain through the use of soil improvers, as well as the possibility that biosolids may represent a cradle for the development of STEC with shuffled virulence genes asset.

All these findings allowed a better comprehension of some of the less known aspects of the epidemiology of STEC infections, such as the involvement of the Stx subtypes in the development of the disease in humans. As a matter of fact, our findings showed that Stx subtypes previously regarded as being adapted to an animal host, as Stx2f, may be able to induce HUS in humans in presence of an efficacious repertoire of determinants that allow establishing a successful colonization of the host intestine.

Our work also provided evidence that the investigation of the threats associated with the re-use of wastes in the framework of the green economy is necessary and has to be faced through an extensive risk assessment exercise. Our work focused on STEC but it was evident from our findings that other threats should also be investigated.

Finally, the investigation of the changes induced in the intestinal microbiota upon infection with a STEC O26:H11 led to the identification of the unbalance in the taxonomic units of the microbiota, which appears to be similar to that observed during infections with other enteric pathogens and described in the literature. This may either be a response of the microbiota to the interaction with the STEC or a general adaptation of the intestinal microflora to the colonization by a pathogen. In any case the knowledge of these dynamics is essential to the comprehension of the pathogenetic mechanisms and paves the way to the development of non-pharmacological approaches to mitigate the burden of intestinal infections. This aspect is particularly interesting for STEC as the antibiotic treatment is not recommended for the treatment of the infections, which mainly relies on prevention and supportive therapy.

In conclusion, the work presented in this thesis aimed at providing knowledge to specifically fill the gaps in our understanding of the complex epidemiology of STEC infections, particularly considering the less knows STEC types and their natural reservoirs, the interactions with the intestinal microbiota in the human hosts and the environmental involvement in the spreading of the infections.

7.5 References

Beutin, L., Bode, L., Ozel, M. and Stephan, R. (1990). Enterohemolysin production is associated with a temperate bacteriophage in Escherichia coli serogroup O26 strains. *J Bacteriol* **172,** 6469-6475.

Brunder, W., Schmidt, H. and Karch, H. (1996). KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic Escherichia coli O157:H7. *Microbiology* **142 (Pt 11),** 3305-3315.

Brunder, W., Schmidt, H. and Karch, H. (1997). EspP, a novel extracellular serine protease of enterohaemorrhagic Escherichia coli O157:H7 cleaves human coagulation factor V. *Mol Microbiol* **24,** 767-778.

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.

de Vos, W.M. and de Vos, E.A. (2012). Role of the intestinal microbiome in health and disease: from correlation to causation. *Nutr Rev* **70 Suppl 1,** S45-56.

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.

Fremaux, B., Prigent-Combaret, C. and Vernozy-Rozand, C. (2008). Long-term survival of Shiga toxin-producing Escherichia coli in cattle effluents and environment: an updated review. *Vet Microbiol* **132,** 1-18.

Friedrich, A.W., Bielaszewska, M., Zhang, W.L., Pulz, M., Kuczius, T., Ammon, A. and Karch, H. (2002). Escherichia coli harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms. *J Infect Dis* **185,** 74-84.

Friesema, I., van der Zwaluw, K., Schuurman, T., Kooistra-Smid, M., Franz, E., van Duynhoven, Y. and van Pelt, W. (2014). Emergence of Escherichia coli encoding Shiga toxin 2f in human Shiga toxin-producing E. coli (STEC) infections in the Netherlands, January 2008 to December 2011. *Euro Surveill* **19,** 26-32.

Gagliardi, J.V. and Karns, J.S. (2002). Persistence of Escherichia coli O157:H7 in soil and on plant roots. *Environ Microbiol* **4,** 89-96.

Garmendia, J., Frankel, G. and Crepin, V.F. (2005). Enteropathogenic and enterohemorrhagic Escherichia coli infections: translocation, translocation, translocation. *Infect Immun* **73,** 2573-2585.

Henderson, H. (2008). Direct and indirect zoonotic transmission of Shiga toxin-producing Escherichia coli. *J Am Vet Med Assoc* **232,** 848-859.

Hugon, P., Lagier, J.C., Colson, P., Bittar, F. and Raoult, D. (2017). Repertoire of human gut microbes. *Microb Pathog* **106,** 103-112.

Ihekweazu, C., Carroll, K., Adak, B., Smith, G., Pritchard, G.C., Gillespie, I.A*., et al.* (2012). Large outbreak of verocytotoxin-producing Escherichia coli O157 infection in visitors to a petting farm in South East England, 2009. *Epidemiol Infect* **140,** 1400-1413.

Jerse, A.E., Yu, J., Tall, B.D. and Kaper, J.B. (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.

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.

Karch, H. (2001). The role of virulence factors in enterohemorrhagic Escherichia coli (EHEC)--associated hemolytic-uremic syndrome. *Semin Thromb Hemost* **27,** 207-213.

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.

Lagier, J.-C. (2016). Gut microbiota and *Clostridium difficile* infections. *Human Microbiome Journal* **2,** 10-14.

Ludden, C., Reuter, S., Judge, K., Gouliouris, T., Blane, B., Coll, F*., et al.* (2017). Sharing of carbapenemase-encoding plasmids between Enterobacteriaceae in UK sewage uncovered by MinION sequencing. *Microb Genom* **3,** e000114.

Mezzanotte, V., Antonelli, M., Citterio, S. and Nurizzo, C. (2007). Wastewater disinfection alternatives: chlorine, ozone, peracetic acid, and UV light. *Water Environ Res* **79,** 2373-2379.

Newton, H.J., Sloan, J., Bulach, D.M., Seemann, T., Allison, C.C., Tauschek, M.*, et al.* (2009). Shiga toxin-producing Escherichia coli strains negative for locus of enterocyte effacement. *Emerg Infect Dis* **15,** 372-380.

Nielsen, P.H. (2017). Microbial biotechnology and circular economy in wastewater treatment. *Microb Biotechnol* **10,** 1102-1105.

Nordmann, P., Naas, T. and Poirel, L. (2011). Global spread of Carbapenemase-producing Enterobacteriaceae. *Emerg Infect Dis* **17,** 1791-1798.

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.

O'Callaghan, A. and van Sinderen, D. (2016). Bifidobacteria and Their Role as Members of the Human Gut Microbiota. *Front Microbiol* **7,** 925.

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.

Persson, S., Olsen, K.E., Ethelberg, S. and Scheutz, F. (2007). Subtyping method for Escherichia coli shiga toxin (verocytotoxin) 2 variants and correlations to clinical manifestations. *J Clin Microbiol* **45,** 2020-2024.

Prager, R., Fruth, A., Siewert, U., Strutz, U. and Tschäpe, H. (2009). Escherichia coli encoding Shiga toxin 2f as an emerging human pathogen. *Int J Med Microbiol* **299,** 343-353.

Rolhion, N. and Chassaing, B. (2016). When pathogenic bacteria meet the intestinal microbiota. *Philos Trans R Soc Lond B Biol Sci* **371**.

Saveyn, H. and Eder, P. (2014) End-of-waste criteria for biodegradable waste subjected to biological treatment (compost & digestate): Technical proposals. EU Commission, IPTS Sevilla (Spain).

Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P.*, et al.* (2009). Organised genome dynamics in the Escherichia coli species results in highly diverse adaptive paths. *PLoS Genet* **5,** e1000344.

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.

Tzouvelekis, L.S., Markogiannakis, A., Psichogiou, M., Tassios, P.T. and Daikos, G.L. (2012). Carbapenemases in Klebsiella pneumoniae and other Enterobacteriaceae: an evolving crisis of global dimensions. *Clin Microbiol Rev* **25,** 682-707.

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.

Appendix

List of Tables and Figures:

EXTERNAL EVALUATION REPORT

Candidate (Name, Surname): Federica Gigliucci

PhD thesis entitled: Shiga-toxin-producing Escherichia coli (STEC): Evolution of the different pathogenic clones and study of their interactions with the microbiota in the intestine and the environment

Reviewer (Name, Surname): Prof. Dr. Med. Alexander W. Friedrich

Institution & Address: Medical Microbiology and Infection Prevention, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 97300RB Groningen, The Netherlands

E-mail: alex.friedrich@umcg.nl

A. Specific questions

Query	Answer (yes/no)
The PhD thesis provides an original contribution to current knowledge in the field (yes/no)	Yes
Aims of the PhD thesis are clearly identified (yes/no)	Yes
The PhD thesis denotes adequate knowledge of the literature in the field (yes/no)	Yes
The PhD thesis denotes that the candidate has sufficient methodological and technical skills to conduct the research (yes/no)	Yes
Results are clearly illustrated (yes/no)	Yes
Results are properly discussed (yes/no)	Yes
The style of text and figures is satisfactory (yes/no)	Yes
References are exhaustive and updated (yes/no)	Yes

B. Comments on the acceptability of the PhD thesis for the final dissertation (free text)

The doctoral thesis presented by Federica Gigliucci focuses on the ecology of pathogenic E. coli bacteria and presents for the first time new insights on the zoonotic power of the whole group of Shiga toxin-producing E. coli. The thesis uses modern technology, comprising whole genome analysis by performing next-gen sequencing of isolates from humans and environmental habitat in order to understand ecological behavior and transmission dynamics.

Prof. Paolo Visca - Dipartimento di Scienze - Università ROMA TRE, viale Marconi 446 - 00146 - Roma, tel+39-06-57336347, fax +39-06-57336321; cellulare +39-.329.0570955; E-mail: paolo.visca@.uniroma3.it

Different to the existing knowledge of quite separated ecological niches, were STEC of different epidemiological background can be contributed specifically to one habitat or the human, Federica Gigliucci's work elucidates a totally different situation today. The borders between the habitats seem to be much less existant any more. STEC that 10 years ago were believed to belong to a specific habitat (e.g. stx2f) are nowadays found in different human populations and diverse environmental niches. It conforms global development of enhanced cross-habitat transmission and new forms of contamination for animals and especially humans. Mrs. Gigliucci even proposes a possible anthropogenic vehicle that can contribute to enhanced cross-species and cross-habitat transmission, the biosolids. Biosolids are an important economic component in modern agriculture and agribusiness. The possibility that biosolids can contribute to a higher epidemic power of different human-pathogenic E. coli is of great concern for public health and shows the importance of OneHealth approaches in research and microbiological risk assessment in public health microbiology.

I appreciate the scientific contribution combining modern microbiological techniques with public health microbiology for a clinically highly relevant bacterial subspecies and consider the thesis as fully acceptable for a doctorate according to international scientific standards.

Kind regards,

Prof. Dr. Alex W. Friedrich
Chair and head of department
Medical Microbiology and Infection Prevention
University Medical Center Groningen
The Netherlands
www.mmb-umcg.nl

Prof. Paolo Visca - Dipartimento di Scienze - Università ROMA TRE, viale Marconi 446 - 00146 - Roma, tel+39-06-57336347,
fax +39-06-57336321; cellulare +39-.329.0570955; E-mail: paolo.visca@.uniroma3.it

EXTERNAL EVALUATION REPORT

Candidate (Name, Surname): **Gigliucci Federica**

PhD thesis entitled:
Shiga-toxin-producing Escherichia coli (STEC): Evolution of the different pathogenic clones and study of their interactions with the microbiota in the intestine and the environment

Reviewer (Name, Surname): **Fach patrick**

Institution & Address:
Agence Nationale de Sécurité Sanitaire (anses)
Laboratoire de sécurité des aliments de Maisons-Alfort
14, rue Pierre et Marie Curie
94701 Maisons-Alfort, France

E-mail: **patrick.fach@anses.fr**

A. Specific questions

Query	Answer (yes/no)
The PhD thesis provides an original contribution to current knowledge in the field (yes/no)	**Yes**
Aims of the PhD thesis are clearly identified (yes/no)	**Yes**
The PhD thesis denotes adequate knowledge of the literature in the field (yes/no)	**Yes**
The PhD thesis denotes that the candidate has sufficient methodological and technical skills to conduct the research (yes/no)	**Yes**
Results are clearly illustrated (yes/no)	**Yes**
Results are properly discussed (yes/no)	**Yes**
The style of text and figures is satisfactory (yes/no)	**Yes**
References are exhaustive and updated (yes/no)	**Yes**

B. Comments on the acceptability of the PhD thesis for the final dissertation (free text)

Mrs Federica Gigliucci has produced in the frame of her PhD thesis a document entitled "**Shiga-toxin-producing Escherichia coli (STEC): Evolution of the different pathogenic clones and study of their interactions with the microbiota in the intestine and the environment**". The document presents an impressive amount of data from different scientific studies on *E. coli*. These studies are linked together by the overall subject of identifying novel or neglected reservoirs and transmission pathways for STEC infections and investigating the relationships of STEC with the microbial communities, once they establish in the human

Prof. Paolo Visca - Dipartimento di Scienze - Università ROMA TRE, viale Marconi 446 - 00146 - Roma, tel+39-06-57336347,
fax +39-06-57336321; cellulare +39-.329.0570955; E-mail: paolo.visca@.uniroma3.it

intestine upon infection. The document provided by Federica Gigliucci has no deficiencies at the scientific level. The work is presented in the form of articles which are already published or submitted for publication in international scientific journals. I have appreciated in particular the coherence in structure of the discussion and the articulation of its parts, proper usage of pertinent documentation and rigor in the argumentation.

This document of 100 pages is separated in different chapters dealing mainly with 1) Stx2f producing-E. coli that represent a serious problem in public health; 2) "Bio-waste" as a new possible transmission pathway of pathogenic E. coli and 3) Characterization of human intestinal microbiota in presence of STEC infections. Furthermore, a thorough introduction presents the current knowledge on subjects relevant for the PhD work and a final discussion covers all studies of the thesis. The thesis includes four articles with the doctorate candidate as first author in two of them (those related to metagenomics). Three of the articles are published in scientific journals and have therefore already undergone peer review. According to international standards for a PhD thesis, this is a high number of scientific articles. The research is original – and part of it is innovative- and has major relevance in its field. Different methods are employed mostly WGS and metagenomics. The employed NGS methods and bioinformatics tools are appropriate and well suited for these studies. The obtained data are generally well analyzed and interpreted with comparison to the relevant previously published literature.

The thesis presents in particular interesting results that are applicable in a new 'molecular risk assessment approach'. Results suggested that Stx2f-producing STEC can cause severe human disease, as HUS, in presence of a complete repertoire of virulence-associated determinants. Pigeons should thus be considered reservoir of highly pathogenic STEC strains, or of stx2f-phages, which can be released in the environment and in turn lysogenize E. coli strains with a virulence genes background associated with HUS, representing a serious risk for public health.

Metagenomics used in this work provided evidences that biosolids may act as a new environmental vehicle of important biological threats for human health. The presence of many genomics traits related with different diarrheagenic E. coli pathogroups was detected, including STEC, confirming previous results obtained for the same samples by Real Time PCR. The holistic analysis conducted in this study through the shotgun metagenomics also allowed to identify the presence in the BSO samples of gene sequences associated with the resistance to the main classes of antibiotics, as Fluoroquinolones, Beta-lactams, Streptothricin, Fosfomycin, Vancomycin and Methicillin, as well as with the resistance to the heavy metals.

Federica Gigliucci has also developed an enhanced metagenomic approach to examine changes in the microbial composition of the intestinal microbiota in faecal samples from patients suffering from STEC infection compared to healthy and healed controls, collected during an outbreak caused by a STEC O26:H11. In this work which is submitted for publication she corroborates previous published results showing that in general, faeces collected from infected patients show a lower intestinal abundance of the beneficial microorganisms Bifidobacterium and Clostridiales spp. with respect to those from the controls group, where those microorganisms predominate. These findings suggest the occurrence of a possible competition between beneficial microorganisms and STEC strains in human intestine and open the way to the study of measures to mitigate the impact of STEC infections and their progression towards the most severe forms.

Generally, the thesis discussion gives good summary of results and puts the results into perspective in comparison to the literature as well as makes the relevant connections between articles in the thesis.

This is a very good thesis which includes in particular a significant and impressive amount of work in metagenomics. The docotorate candidate has demonstrated a high level of knowledge and understanding of the subject. The document is well written and clearly presented.

Patrick FACH
Coordinateur de la Plateforme IdentyPath
Agence nationale de sécurité sanitaire de l'alimention,
de l'environnement et du travail (Anses)
Laboratoire de sécurité des Aliments de Maisons-Alfort
14, rue Pierre et Marie Curie,
94701 Maisons-Alfort Cedex, France
Tél : +33 (0)1 4977 2813
Mobile : +33 (0)6 8059 3050
patrick.fach@anses.fr

Prof. Paolo Visca - Dipartimento di Scienze - Università ROMA TRE, viale Marconi 446 - 00146 - Roma, tel+39-06-57336347,
fax +39-06-57336321; cellulare +39-.329.0570955; E-mail: paolo.visca@.uniroma3.it