



C-source metabolic profilings of foodborne Shiga-toxin producing *E. coli* match serogroup differentiations and highlight functional adaptations

Stéphane Kerangart, Benoit Cournoyer, Estelle Loukiadis

► To cite this version:

Stéphane Kerangart, Benoit Cournoyer, Estelle Loukiadis. C-source metabolic profilings of foodborne Shiga-toxin producing *E. coli* match serogroup differentiations and highlight functional adaptations. International Journal of Food Microbiology, 2017, 266, pp.324-336. 10.1016/j.ijfoodmicro.2017.10.018 . hal-02487190

HAL Id: hal-02487190

<https://univ-lyon1.hal.science/hal-02487190>

Submitted on 6 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



C-source metabolic profilings of foodborne Shiga-toxin producing *E. coli* match serogroup differentiations and highlight functional adaptations

Stéphane Kerangart^a, Benoit Cournoyer^{a,*}, Estelle Loukiadis^{a,b,1}

^a University of Lyon, Research group on “Bacterial Opportunistic Pathogens and Environment”, UMR Ecologie Microbienne, CNRS 5557, INRA 1418, University Lyon 1, VetAgro Sup, 69280 Marcy L'Etoile, France

^b University of Lyon, VetAgro Sup, LMAP, National Reference Laboratory for *Escherichia coli* (including VTEC), Marcy L'Etoile, France

ARTICLE INFO

Keywords:

STEC
Metabolic profilings
Carbon sources
INT-reduction assays
O157
O26
O104:H4

ABSTRACT

The tropism of pathogenic STEC for foodstuffs and cattle reservoir is related to functional specializations. An investigation of C-source utilization patterns among and between STEC serogroups was performed using omnilog phenotypic microarrays (OM). OM functional groupings were compared with STEC phylogroups, seropathotypes, EFSA's molecular risk assessment groups and serogroups. OM INT reduction activities of 37 STEC strains growing on 190 C-substrates were compared. Each strain had its own specific C-utilization profile but 23% of the substrates was used by all strains, 47% by none, and 30% was variably metabolized. Galactose, mannose, *N*-acetyl-glucosamine (GlcNAc), and *N*-acetyl neuraminic acid (Neu5Ac) found in the mucus layer of the bovine small intestine were metabolized by all strains. The 56 most informative substrates divided the C-utilization patterns (CP) into three clusters with: (A) harboring all O157 and O145 strains; (B) all O26 strains, and (C) strains of the other serogroups. Significant correlations between INT reduction values of pair of strains per CP group supported these differentiations. CP of group A and B strains were respectively defective in the use of galactonic acid- γ -lactone and rhamnose. Most CP group C strains grew with *L*-lyxose. Adjusted Wallace coefficients analyses of the datasets indicated high probabilities for the prediction of the use of glycolic acid, β -hydroxybutyric acid, *L*-lyxose and *D*-galactonic acid- γ -lactone and 5-keto-*D*-gluconic acid by a serogroup. The use of a C-substrate could be predicted from the classification of a strain into a phylogroup or seropathotype. Significantly lower numbers of C-substrates were used by seropathotype A strains like O157 ones. Improvements of STEC identification keys were proposed using the most discriminant C-substrates found in this study.

1. Introduction

Shiga toxin-producing *Escherichia coli* (STEC) strains are major worldwide foodborne pathogens, causing severe and sometimes lethal human infections, such as hemolytic uremic syndrome (HUS) cases, especially among young children (Croxen et al., 2013; Kaper et al., 2004). They are notably responsible of outbreaks and sporadic cases of acute illnesses (reaching each year 2,801,000 cases), and cause annually 3890 HUS cases and 230 deaths (Majowicz et al., 2016). The key virulence trait of STEC is the synthesis of Shiga toxins (Stx) but additional virulence factors boosting their pathogenicity have been described (Karmali, 2009). Most typical pathogenic strains harbor a LEE (locus of enterocyte effacement) involved in the intimate adherence to the host-cell membrane (Franzin et al., 2015). Recent outbreaks, such as the one involving STEC O104:H4 strains lacking the LEE, led to the observation of additional virulence properties (Beutin and Fach, 2014;

Navarro-Garcia, 2014). STEC virulence thus appears multi-factorial, and several factors remain to be identified (Karmali et al., 2010).

More than 380 STEC serotypes have been reported in humans so far but only ten of them have been consistently associated with severe cases and foodborne outbreaks (Caprioli et al., 2005; Mathusa et al., 2010). To improve the assessment of clinical and public health risks associated with STEC, Karmali et al. (2003) proposed a seropathotype classification (using a gradient from seropathotype A – high risk- to seropathotypes D and E – minimal risk). This classification considers the prevalence of STEC serotypes and their association with human outbreaks, including the severity of the clinical outcomes (Karmali et al., 2003). Serotypes O157:H7 and O157:NM have been classified in the seropathotype A, those of serotypes O26:H11; O103:H2; O111:NM; O121:H19 and O145:NM into seropathotype B, and those of O91:H21 and O113:H21 into seropathotype C. Recently, a new scheme for STEC categorization according to potential risk for human health was

* Corresponding author.

E-mail address: benoit.cournoyer@vetagro-sup.fr (B. Cournoyer).

¹ Contributed equally to this work.

Table 1Serotype, seropathotype, molecular risk assessment and phylogenetic groups, and origin of the 37 O157 and non-O157 *E. coli* strains used in this study.

Serogroup	Strain	Serotype	Pathotype	STEC seropathotype ^a	STEC molecular risk assessment group ^b	Phylogenetic group ^c	Origin	References
O157	620	O157:H7	STEC	A	I	D	Host independent origin (Food-bovine meat)	King et al., 2014
	RD9	O157:H7	STEC	A	I	D	Host independent origin (Food-bovine meat)	King et al., 2009
	540	O157:H7	STEC	A	I	D	Host independent origin (Food-bovine meat)	Kerangart et al., 2016
	FCH6	O157:H7	STEC	A	I	D	Host independent origin (Food-raw goat's milk cheese)	Espié et al., 2006
	EDL 933	O157:H7	STEC	A	I	D	Host gut (Human)	Perna et al., 2001
	Sakai	O157:H7	STEC	A	I	D	Host gut (Human)	Hayashi, 2001
	m622	O157:H7	STEC	A	I	D	Host independent origin (Food-raw milk cheese)	Franz et al., 2011
O26	m638	O157:H7	STEC	A	I	D	Host gut (Human)	Franz et al., 2011
	21765	O26:H11	STEC	B	I	B1	Host independent origin (Food-raw cow milk cheese)	Espié et al., 2006
	2245/98	O26:H11	STEC	B	I	B1	Host gut (Human)	Zimmerhackl et al., 2010
	64/36	O26:H11	AEEC	ND	ND	B1	Host independent origin (Food-raw ewe milk)	Neto et al., 2012
	3073/00	O26:H11	STEC	B	I	B1	Host gut (Human)	Zimmerhackl et al., 2010
	37.40	O26:H11	STEC	B	I	B1	Host independent origin (Food-bovine meat)	Neto et al., 2012
	279/8	O26:H11	STEC	B	I	B1	Host independent origin (Food-raw cow milk cheese)	Kerangart et al., 2016
	11368	O26:H11	STEC	B	I	B1	Host gut (Human)	Ogura et al., 2009
	MB 2659 (PH10)	O26:H11	STEC	B	I	B1	Host gut (Human)	Possé et al., 2007
	PMK5	O103:H2	STEC	B	I	B1	Host gut (Human)	Mariani-Kurkdjian et al., 1993
O103	262.03.18.1	O103:H-	<i>E. coli</i>	ND	ND	B1	Host independent origin (Food-bovine meat)	This study
	MB 2652 (PH3)	O103:H2	STEC	B	I	B1	Host gut (Human)	Possé et al., 2007
	MB 2653 (PH4)	O103:H2	STEC	B	I	B1	Host gut (Human)	Possé et al., 2007
							Host gut (Human)	
O111	CL37	O111:H8	STEC	B	I	ND	Host gut (Human)	Zhang et al., 2002
	J43	O111:H8	STEC	B	I	ND	Host gut (Bovine)	Bibbal et al., 2015
	622-4	O111:H8	STEC	B	I	ND	Host independent origin (Food-raw cow milk cheese)	Kerangart et al., 2016
O145	2513-21	O145:H28	STEC	B	I	D	Host independent origin (Food-raw cow milk cheese)	Kerangart et al., 2016
	991	O145:H28	STEC	B	I	D	Host independent origin (Food-bovine meat)	Kerangart et al., 2016
	MB 2817 (PH88)	O145:H28	STEC	B	I	D	Host gut (Human)	Possé et al., 2007

(continued on next page)

Table 1 (continued)

Serogroup	Strain	Serotype	Pathotype	STEC seropathotype ^a	STEC molecular risk assessment group ^b	Phylogenetic group ^c	Origin	References
O121	12409	O121:H11	<i>E. coli</i>	ND	ND	ND	Host gut (Bovine)	This study
	12652	O121:H19	STEC	B	II	B1	Host gut (Human)	Kerangart et al., 2016
	12805	O121:H19	STEC	B	II	B1	Host gut (Human)	Kerangart et al., 2016
O45	12047	O45:H2	STEC	ND	II	ND	Host gut (Human)	Kerangart et al., 2016
O104	Zucker-O104	O104:H4	STEC	ND	I	ND	Host gut (Human)	King et al., 2012
O91	13199	O91:H21	STEC	C	III	B1	Host independent origin (Food-bovine meat)	Kerangart et al., 2016
	B2F1	O91:H21	STEC	C	III	B1	Host gut (Human)	Ito et al., 1990
	13694	O91:H21	STEC	C	III	B1	Host independent origin (Food-raw cow milk)	Kerangart et al., 2016
O113	13341	O113:H21	STEC	C	III	B1	Host independent origin (Food-raw cow milk)	Kerangart et al., 2016
	14032	O113:H21	STEC	C	III	B1	Host independent origin (Food-bovine meat)	Kerangart et al., 2016
	13137	O113:H4	STEC	D	III	A	Host independent origin (Food-sheep meat)	Kerangart et al., 2016

ND: not done, STEC: Shiga-toxin producing *E. coli*; AEEC: *E. coli* able of attachment and effacement; a: according to Karmali et al. (2003), b: according to EFSA (2013), c: according to Girardeau et al. (2005) and Mora et al. (2012).

proposed (EFSA, 2013) ranging from group I (including STEC belonging to O157, O26, O103, O145, O111 and O104 serogroups, and considered of higher risk), to group III (STEC strains for which there is an uncertainty regarding their ability at causing a disease).

Asymptomatic domestic ruminants have been identified as the main reservoir of highly virulent serogroups of STEC strains, with a global fecal carriage of beef and dairy cattle reaching up to 30% for O157 and 70% for non-O157 strains (Persad and LeJeune, 2014). This carriage explains that most human infections have been related to the consumption of contaminated food of bovine origin such as ground beef, raw milk and related products such as cheese (Feng, 2014). Direct contact with healthy animal carriers and exposure to contaminated environments such as soils, runoff waters, bathing waters and crops could also explain some sporadic cases and outbreaks (Fegan and Gobius, 2012; Saxena et al., 2015). The tropism of pathogenic STEC for the above habitats or niches is likely related to particular metabolic requirements and ecological specializations (Bertin et al., 2013; Farrokh et al., 2013; Pifer and Sperandio, 2014). In fact, carbohydrates of the mucus layer covering enterocytes of the bovine small intestine such as mannose, *N*-acetyl-glucosamine (GlcNAc), *N*-acetyl neuraminic acid (Neu5Ac) and galactose were found to increase O157:H7 EDL933 STEC strain growth rate and favor their development over other species of the microbiota including commensal *E. coli* (Bertin et al., 2013). Maltose, some short-chain oligosaccharides, deoxyribose, and synthesis of glycogen, were also found to play an important role in O157:H7 EDL933 STEC strain host colonization (Jones et al., 2008; Schouler et al., 2009).

STEC metabolic specializations have also been highlighted through small scale growth screening on various media. These screenings led to a use of certain substrates including antimicrobials in the STEC identification schemes. For example, O157:H7 strains were differentiated on the basis of sorbitol fermentation (March and Ratnam, 1986) or resistance towards tellurite (Kerangart et al., 2016). D-saccharic acid, D-serine, D-galactonic acid-γ-lactone, and glycolic acid were found differentiating O157:H7 isolates from commensal *E. coli* e.g. Durso et al., 2004. Raffinose, dulcitol, and rhamnose were also suggested as

discriminatory substrates for this STEC serotype (Ratnam et al., 1988). However, non-O157 strains have been considered highly diversified, and difficult to differentiate on the basis of growth requirements (Bettelheim and Beutin, 2003; Beutin and Fach, 2014; Schmidt et al., 1999) even though carbohydrate utilization patterns (rhamnose, dulcitol, raffinose and arabinose) have been shown to discriminate O26, O103, O111 and O145 strains (Possé et al., 2007). The ecological benefits of such C-source utilization patterns remain to be understood. Evolution of these patterns is likely in line with phylogenetic similarities/divergences, matching gains or losses of certain metabolic pathways (Sabarwal et al., 2011). Still, wide phenotypic variations among the species were reported through inferences from the *E. coli* pan-genome (Rasko et al., 2008; Touchon et al., 2009).

Here, phenotypic traits of foodborne and bovine STEC isolates were further studied with the aim of identifying characters matching their lifestyle or phylogenetic diversifications. STEC C-source utilization patterns (CP) were investigated, and compared with STEC groupings into phylogroups, seropathotypes, EFSA's molecular risk assessment (MRA) groups and serogroups. These analyses were performed on a panel of 37 STEC strains, representative of host-dependent and host-independent groupings. Our results indicated C-source metabolic profiles to be in line with the vertical evolution of STEC strains and to be relevant for identifying recent diversification events among serogroups.

2. Materials and methods

2.1. Bacterial strains

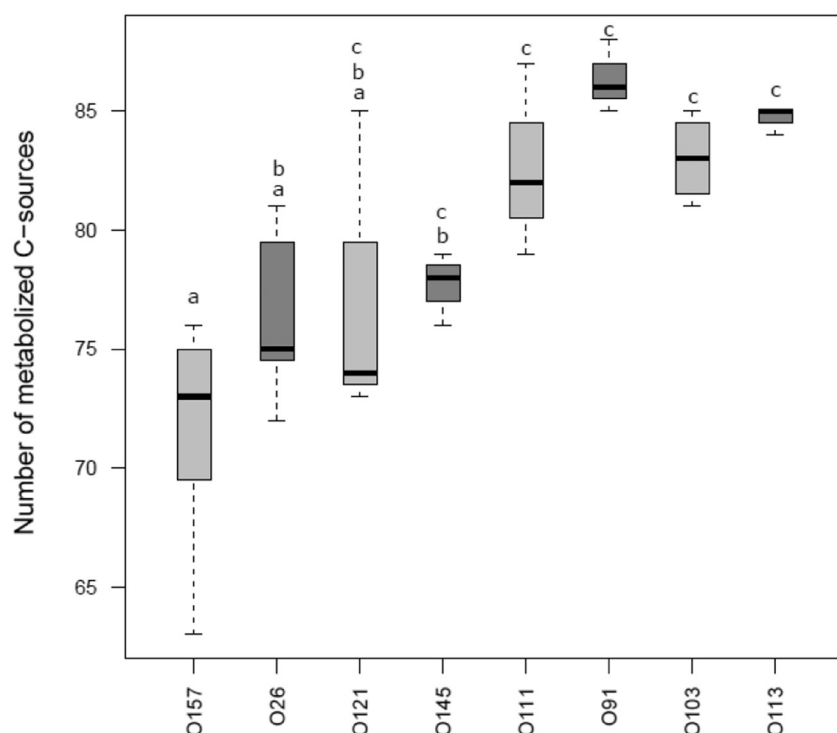
A panel of 37 STEC strains including two deleted of their *stx* genes, one from O103 and one from O121, from the German and French national reference collections (VetAgro Sup, Marcy l'Etoile, France, and BfR, Berlin, Germany), isolated either from the human or bovine gastrointestinal tracts (*n* = 19) or from foodstuffs (*n* = 18), was selected (Table 1). Particular attention was paid to build up a collection representative of the most frequent food matrices involved in human clinical cases (bovine and sheep meat; raw milk cheeses and milk from

Table 2List of informative* C-substrates discriminating STEC lineages by a maximum parsimony analysis¹.

PM1 A04 D-Saccharic acid (3)	PM1 D10 Lactulose (45)	PM2A A06 Dextrin (100)
PM1 A08 L-Proline (7)	PM1 D11 Sucrose (46)	PM2A B01 N-Acetyl-D-Galactosamine (107)
PM1 A09 D-Alanine (8)	PM1 E01 L-Glutamine (48)	PM2A B03β-D-Allose (109)
PM1 A12 Dulcitol (11)	PM1 E02 m-Tartaric acid (49)	PM2A B05 D-Arabinose (111)
PM1 B01 D-Serine (12)	PM1 E07 α-Hydroxybutyric acid (54)	PM2A B09 2-Deoxy-D-Ribose (115)
PM1 B02 D-Sorbitol (13)	PM1 E08 β-Methyl-D-Glucoside (55)	PM2A B12 3-O-β-D-Galactopyranosyl-D-Arabinose (118)
PM1 B03 Glycerol (14)	PM1 F01 Gly-Asp (60)	PM2A C07 β-Methyl-D-Galactoside (125)
PM1 B10 Formic acid (21)	PM1 F07 Propionic acid (66)	PM2A C09 β-Methyl-D-Glucuronic acid (127)
PM1 B12 L-Glutamic acid (23)	PM1 F08 Mucic acid (67)	PM2A D01 D-Raffinose (131)
PM1 C02 D-Galactonic acid-γ-Lactone (25)	PM1 F09 Glycolic acid (68)	PM2A D04 L-Sorbose (134)
PM1 C04 D-Ribose (26)	PM1 F10 Glyoxylic acid (69)	PM2A D06 D-Tagatose (136)
PM1 C06 L-Rhamnose (29)	PM1 G01 Gly-Glu (72)	PM2A E04 D,L-Citramalic acid (146)
PM1 C08 Acetic acid (31)	PM1 G04 L-Threonine (75)	PM2A E08 β-Hydroxybutyric acid (150)
PM1 C10 Maltose (33)	PM1 G07 Acetoacetic acid (78)	PM2A E12 5-Keto-D-Gluconic acid (154)
PM1 D02 D-Aspartic acid (37)	PM1 H01 Gly-Pro (84)	PM2A F01 D-Lactic acid Methyl Ester (155)
PM1 D03 D-Glucosaminic acid (38)	PM1 H02 p-Hydroxyphenyl Acetic acid (85)	PM2A F03 Melibionidic acid (157)
PM1 D06 α-Ketoglutaric acid (41)	PM1 H03 m-Hydroxyphenyl Acetic acid (86)	PM2A G02 L-Alaninamide (168)
PM1 D07 α-Ketobutyric acid (42)	PM1 H06 L-Lyxose (89)	PM2A H09 Dihydroxyacetone (187)
PM1 D08 α-Methyl-D-Galactoside (43)	PM1 H09 L-Galactonic acid-γ-Lactone (92)	

1: numbers in brackets match those on the tree branches of Fig. 4; *substrates in bold are highly discriminant and were used in the partition analysis of Fig. 6.

*: non-informative substrates were used by all strains (100%) or none.

**Fig. 1.** Boxplot representation of variations in the number of C-substrates metabolized per STEC serotype. *E. coli* INT-reduction activities were monitored over a 48 h incubation period. 190 C-substrates were tested. Each box indicates the variations between the first and third quarter of the obtained values, and the median value (thick black line). Non-parametric Mann-Whitney Wilcoxon statistical tests were performed. A same letter code indicates an absence of significant difference between the data sets ($p > 0.05$), and a distinct letter code indicates significant differences ($p < 0.05$). The strains used in this analysis are described in Table 1.

cow, goat and ewe). Seven of these STEC strains have been associated with food outbreaks linked to ingestion of contaminated meat (King et al., 2014, 2009; Perna et al., 2001), raw milk cheeses (Espíe et al., 2006) or sprouts (Hayashi, 2001; King et al., 2012). Strains isolated from human or bovine fecal matters were from previous reports (Bibbal et al., 2015; Grad et al., 2012; Mariani-Kurkdjian et al., 1993; Perna et al., 2001; Vernozy-Rozand et al., 2004; Zhang et al., 2002; Zimmerhackl et al., 2010). All human strains were isolated from HUS or HC cases. Food strains were isolated according to the ISO 16654 or ISO TS 13136:2012, and have been described elsewhere (Espíe et al., 2006; Franz et al., 2011; Hayashi, 2001; Kerangart et al., 2016; King et al., 2014, 2012, 2009; Perna et al., 2001).

STEC strains were selected to be representative of the ten most clinically relevant serotypes: O157 (n = 8), O26 (n = 8), O103 (n = 4), O113 (n = 3), O121 (n = 3), O145 (n = 3), O111 (n = 3), O91 (n = 3), O45 (n = 1) and O104 (n = 1). Phylogenetic groups of

the strains studied were inferred from Girardeau et al. (2005) and Mora et al. (2012). To consider the public health risks associated with the selected strains, their assignment to seropathotypes and molecular risk assessment groups were performed according to the actual guidelines (Karmali et al., 2003; EFSA, 2013).

All strains were kept at -80°C using Microbank™ beads (Pro-Lab Diagnostics, Neston, UK). Prior performing the growth tests, beads were transferred into a broth of Brain Heart Infusion (BHI) (bioMérieux, Marcy l'Etoile, France) and incubated overnight at 37°C under aerobic and static conditions. Broths were then streaked onto Plate Count Agar (PCA) (bioMérieux, Marcy l'Etoile, France) to produce isolated colonies. Six strains previously used in C-source metabolic profilings were included in this study: O26:H11 strain MB2659 (PH10); O103:H2 strains MB2652 (PH3) and MB2653 (PH4); O145:H28 strain MB2817 (PH88) (Possé et al., 2007) and O157:H7 strains m622 and m638 (Franz et al., 2011).

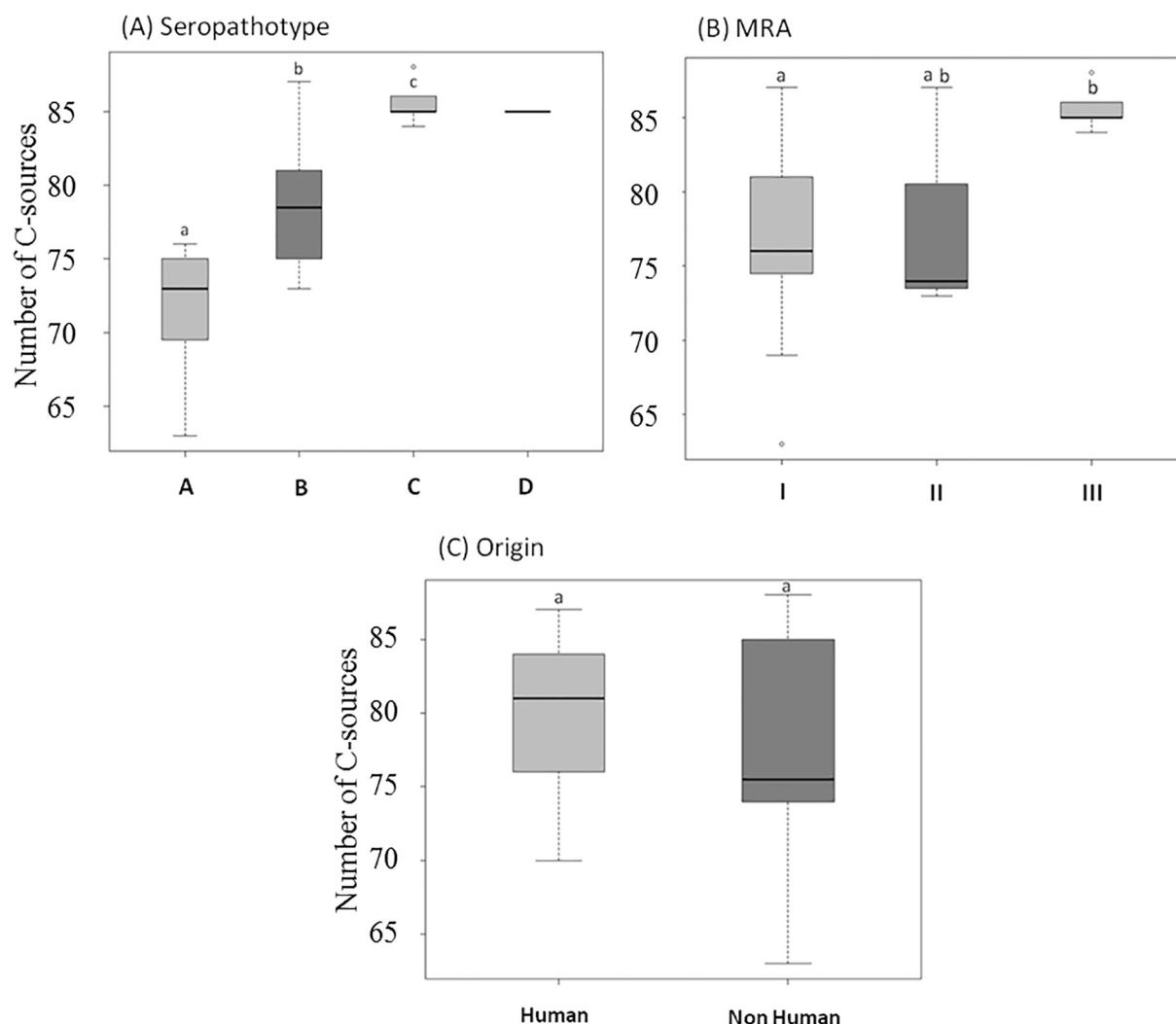


Fig. 2. Boxplot representation of variations in the number of C-substrates metabolized per (A) seropathotype, (B) molecular risk assessment — MRA group, and (C) according to the origin (human or not) of the strains. *E. coli* INT-reduction activities were monitored over a 48 h incubation period. Each box indicates the variations between the first and third quarter of the obtained values, and the median value (thick black line). Non-parametric Mann-Whitney Wilcoxon statistical tests were performed. A same letter code indicates an absence of significant difference between the data sets ($p > 0.05$), and a distinct letter code indicates significant differences ($p < 0.05$). Strains used in these analyses are described in Table 1.

2.2. Preparation of phenotypic microarray (PM) microplates

C-substrate utilization patterns were analysed by phenotypic microarrays (PM) (Biolog, Hayward, CA). The PM technology is based on the irreversible reduction of INT (tetrazolium violet) to formazan (purple color) by NADH (cell respiratory chain). This reaction can be used to monitor growth development. The Biolog PM1 and PM2 microplates were used, and allowed testing of 190 C-substrates (carbohydrates, amino acids, nucleosides, alcohols). Each plate included a negative control. Inoculation of PM plates and preparation of the different inoculating fluids were done according to Biolog recommendations for *E. coli*. Briefly, an *E. coli* colony from a 24 h-old plate kept at 37 °C was picked and re-suspended at a light transmittance of 85%, in the inoculating fluid (IF-0; supplied by BIOLOG and containing the INT dye). Each well of the PM plate was inoculated with 100 µL of cell suspension. The plate was then transferred at 37 °C in the Omnilog reader. INT-reduction datasets were generated by monitoring formazan formation over time at 15 min time-steps over a period of 48 h. Charge-coupled device (CCD) images were used to measure formazan light emissions which were transformed into Omnilog arbitrary units. Datasets were verified using the parametric and kinetic Omnilog PM software.

2.3. Kinetics of STEC INT-reduction activities over time

Omnilog (OM) readings representative of formazan light emissions were plotted over time, and used to compute the area under the curves. These values were extracted and transformed either into binary characters (0/1; 0 = no significant OM values and 1 = significant values, see below) or normalized in a statistical frame that was previously tested (Vehkala et al., 2015). In order to avoid experimental biases and detect the most significant trends in the datasets, area under the curves were computed from 48 h recorded values. Some strains showed a significant lag phase for certain substrates at the beginning of their incubation period. Readings over 48 h prevented a misinterpretation of the growth patterns. Transformation of OM values area under the curves into 0/1 characters was done by computing ratios between the *E. coli* datasets and the background control values (min = 9039; max = 25,184; median = 16,629.5). These ratios were then interpreted as being significant (scored 1) or non-significant (scored 0) as previously done by Franz et al. (2011). A ratio of < 2.0 was considered non-significant. Ratios close to 2 were verified manually. Their OM values plotted over time were analysed in order to detect significant shifts indicative of active growth. This led to the identification of a few cases ($n = 22$), where ratios between 1.8 and 1.9 had to be re-classified

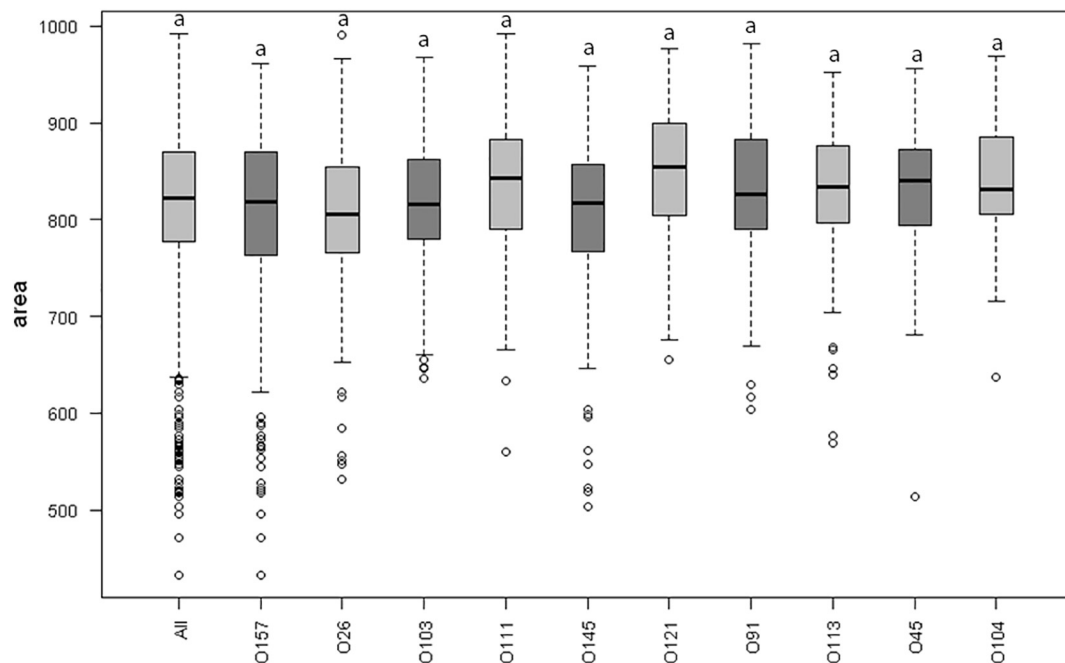


Fig. 3. Boxplot representation of area under the curves computed from shiga-toxin producing *E. coli* INT-reduction activities over a 48 h incubation period, and using the 44 C-sources metabolized by all strains of the tested serogroups. Each box indicates the variations between the first and third quarter of the dataset, and the median value (thick black line). Outliers are indicated by circles. The outliers of the first boxplot ("All" strains) are shown in Table S4. A same letter code indicates the absence of significant difference between the datasets ($p > 0.05$). Computed area under the curves were normalized and transformed (see Materials and methods). Strains used per serogroup are indicated in Table 1.

into the category of significantly active cells. These exceptions were thus given a score of "1". A complete overview of the computed ratios is available in Tables S1 and S2.

2.4. Analysis of C-substrate utilization patterns and statistical tests

The matrix of binary characters was analysed using a parsimony analysis of the dissimilarities observed between strains. The PHYLIP package (Felsenstein, 1985) was used to compute the most parsimonious (MP) clusters. The Wagner parsimony method, allowing changes in both the $0 \rightarrow 1$ and $1 \rightarrow 0$ directions was used. A consensus cladogram of the most frequently recovered MP clusters was built with the Consense programme. This tree was visualized by the MacClade graphic tool (Maddison, 2005). Changes in C-source utilization patterns (CP) between strains were traced using MacClade, and led to the definition of clusters of profiles named CP_A, CP_B, and CP_C.

For statistical analyses, area under the curves computed from OM values were normalized against the highest value obtained (value = 64,749) in the frame of this study. These normalized datasets were then transformed by an arbitrary multiplication of 1000. These transformations did not impact the statistical tests but allowed a better graphical representation of the datasets. Boxplot representations and statistical tests were all performed using the R package (version 3.0.1, R Core Team., 2013). The adjusted Wallace coefficients analyses enabling comparisons of partitions between STEC classifications and the use or not of a C-substrate were computed according to Severiano et al. (2011) using UMMI web site at Universidade de Lisboa (<http://www.comparingpartitions.info/>).

The statistical analyses performed in this study involved the use of $n = 3$ to $n = 25$ strains (or independent omnilog phenotypic microarray (OM) datasets) per category or group of datasets. To investigate the relations between numbers of C-substrates used per serogroup, those with a single strain (O45 and O104) were not considered. To test the variations in the number of C-substrates metabolized per STEC molecular risk assessment (MRA) group, $n = 25$ strains for group I, $n = 3$ strains for group II, and $n = 6$ for group III, were used. To test the variations in the number of C-substrates metabolized according to

the origin of the strains; 17 STEC strains of human origin and 20 strains of food and animal origins were used. The correlation tests between C-source utilization patterns (using relative INT reduction values) per pair of strains involved the use of $n = 56$ values per strain. For the adjusted Wallace statistical tests performed on the binary OM datasets based on the use of the twelve most discriminant substrates found in this study, the following criteria were applied: (a) $n = 11$ strains from the D phylogroup and $n = 19$ for the B1 phylogroup; (b) $n = 8$ strains from the A STEC seropathotype, $n = 18$ for the B seropathotype, and $n = 5$ for the C one; (c) $n = 3$ to 8 strains per serogroup; (d) $n = 10$ strains for CP_A, $n = 10$ strains for CP_B, and $n = 17$ strains for CP_C. The Adjusted Wallace coefficients gave the probability associated with the prediction of the use of one of the tested substrates by a strain according to its classification into a CP group or phylogroup or serogroup or seropathotype (according to the tested combination).

3. Results and discussion

3.1. Conserved C-source utilization patterns

Previous analyses of STEC C-source utilization patterns were performed on the PM1 plate (95 substrates) but were limited to the analysis of O157:H7 isolates and comparisons with commensal *E. coli* e.g. Durso et al. (2004) and Franz et al. (2011). These investigations led to the identification of C-substrates that could be indicative of O157:H7 functional diversifications but comparisons with other clinically-relevant serogroups had not been performed. These previous studies showed 28 of the 95 substrates of Omnilog PM1 plate to be used by all tested *E. coli* strains ($n = 122$) including 81 O157:H7 isolates (Durso et al., 2004). In our investigation, 37 strains from ten serogroups, O26, O45, O91, O103, O104, O111, O113, O121, O145, and O157, were used. A ratio analysis of INT reduction activities over control (no substrate) values was used to identify discriminant substrates (Ratios ≥ 2 ; Tables S1 and S2), and showed (1) 44 (23%) C-sources over the 190 substrates of Omnilog PM1 and PM2 plates to be metabolized by all of the selected strains (Table S3), (2) 90 C-substrates (47%) by none (Table S3), and (3) 56 (30%) to be variably metabolized (Table 2).

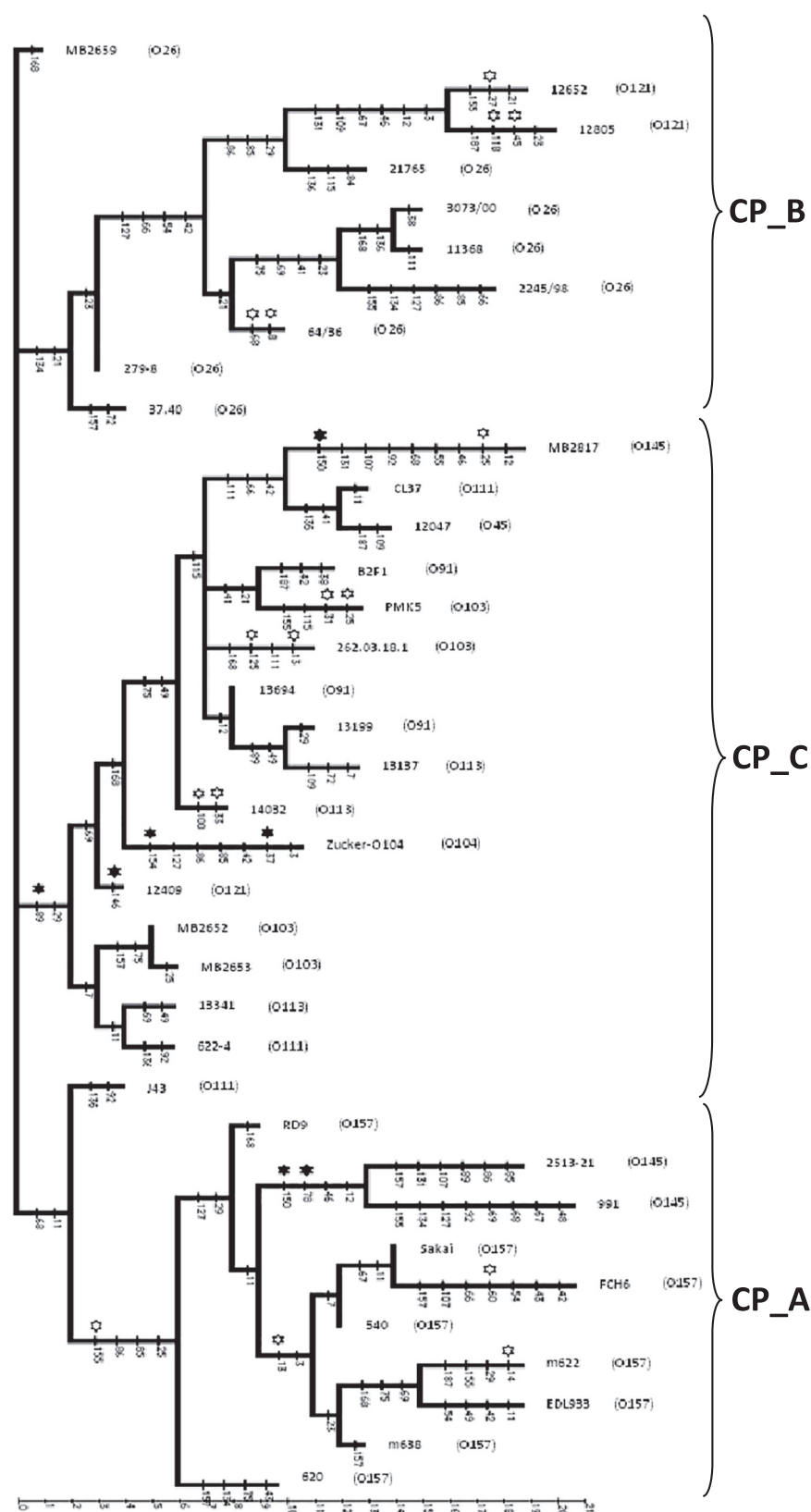


Fig. 4. Cladogram illustrating the relationships between C-source utilization patterns of STEC strains. Pairwise differences between profiles were inferred by a maximum parsimony (MP) analysis of INT-reduction values transformed into binary characters, with 0 indicating a non-significant value and 1 a significant one. See [Materials and methods](#) for the computation. The 56 informative characters are indicated on each branch of the MP tree, and their matching substrate (number on the line) is indicated in Tables 2 and S5. C-substrates discriminating a particular lineage among the full MP tree are indicated by a black star if metabolized or a white star if not metabolized. These C-utilization patterns (CP) were divided into three main clusters, CP_A, CP_B, and CP_C that were supported by the statistical tests shown in [Figs. 5 and 6](#).

These latter substrates were considered informative and useful for discriminating strains. A boxplot analysis of the number of C-substrates used by a serogroup (those with a $n > 3$ strains, [Table 1](#)) was performed ([Fig. 1](#)). Significantly lower numbers of C-substrates were used by O157 strains than O145, O111, O91, O103 or O113 ones. However,

no significant differences between numbers of substrates used by O157, O26 and O121 strains were found ([Fig. 1](#)). Grouping the results (number of substrates) according to the seropathotype classification confirmed this trend. Seropathotype A (strictly O157:H7 strains, $n = 8$) showed a use of a lower number of substrates, while seropathotype B

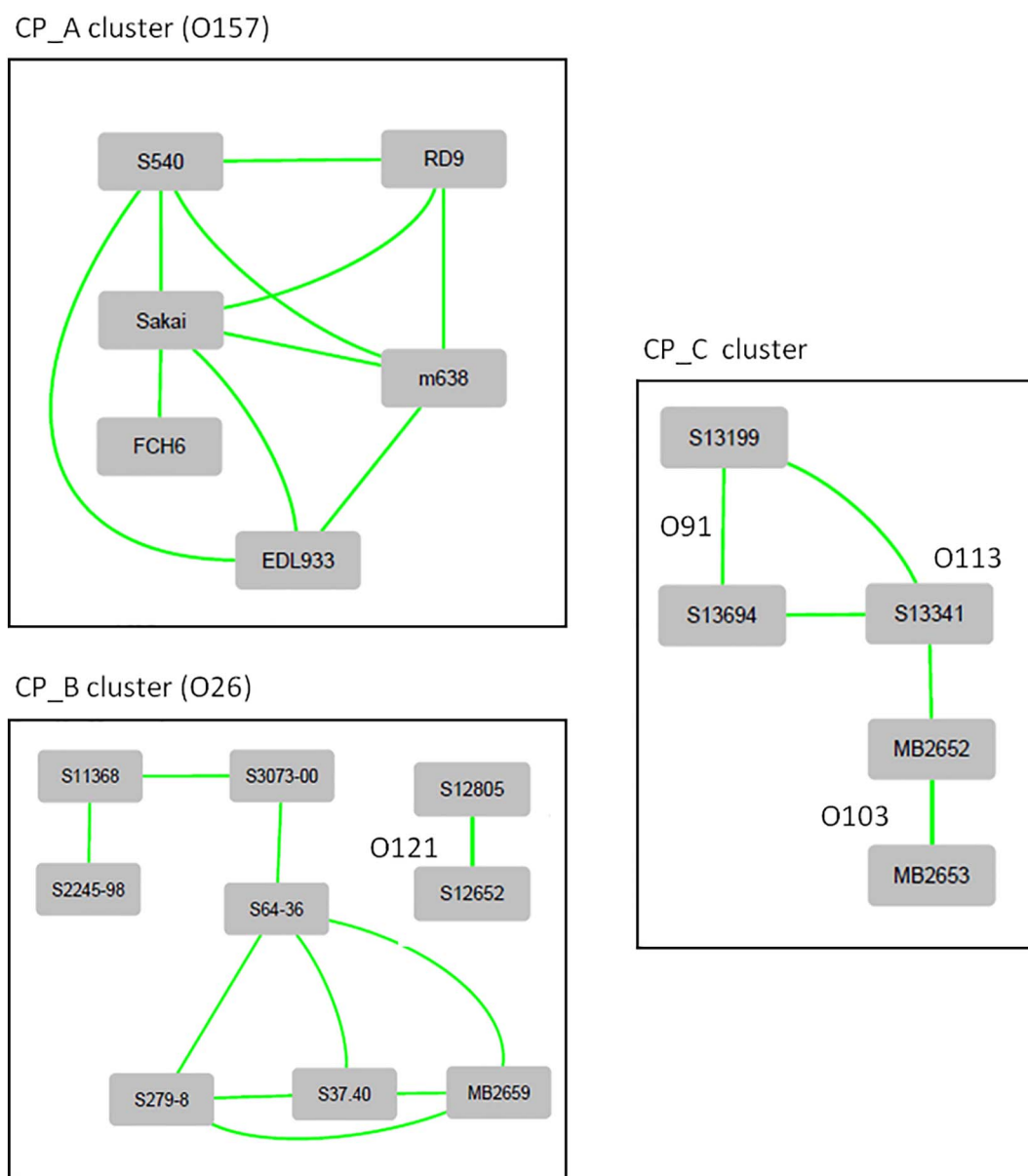


Fig. 5. Illustration of the relatedness between STEC C-source OM profiles through Spearman correlation tests. Area under the curves representative of INT-reduction activities over a 48 h incubation period were compared for all 56 informative C-sources presented in Table 2. Positive and significant correlations (green lines) were observed between OM profiles of strains of the CP_A, the CP_B, and the CP_C clusters of Fig. 4. Area under the curves were normalized and transformed (see Materials and methods) prior performing these tests. Strains are described in Table 1. A green line indicates a positive correlation at $P < 0.05$ and $R^2 > 0.6$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

($n = 18$ strains) and C ($n = 5$ strains) used higher numbers of substrates (Fig. 2A; see Table 1 for the strains used per seropathotype). These results suggest functional specialization for the most hazardous STEC strains which impacted their ability at using a large panel of C-substrates. Grouping strains according to their MRA (molecular risk assessment) classification confirmed a clear differentiation between group III strains ($n = 6$ strains) and the others (from groups I ($n = 25$ strains) and II ($n = 3$ strains) (Fig. 2B). Comparisons between human ($n = 17$ strains) and non-human ($n = 20$ strains) origins showed no significant differences (Fig. 2C).

A global analysis of the INT-reduction activities from substrates ($n = 44$) used by all strains ($n = 37$) was performed (Fig. 3). Some of these substrates were used with variable efficacy suggesting some fine tunings in the regulation of the metabolic pathways that were involved. The measured areas under the INT reduction curves using these substrates showed some significant differences when performing

comparisons per serogroup (Table S4). These variations confirmed likely functional subgroupings among a serogroup. This was previously observed by Franz et al. (2011) for O157 isolates, and sub-groups were related to an adaptation for manure. In fact, 44 values outside the range (outliers) of the lower inner fence of some boxplots appeared to have impacted the statistical tests, and suggested significant differences in glycolysis ($n = 34/44$) and TCA ($n = 9/44$) pathways of some strains among a same serogroup. Most significant differences were related to the less significant growth yields obtained from O157 isolates while using D-glucose (outliers = strains 620, RD9, 540 and FCH6), bromo-succinic acid (outliers = strains FCH6, RD9, 540 and Sakai) and Glucuronamide (outliers = strains 620, RD9, 540, FCH6, EDL933 and Sakai) as C-sources, and from O26 strains 37.40, 11368, and 3073/00 while using Glucuronamide (Table S4). Among these substrates used by all strains, carbohydrates of the mucus layer of the bovine small intestine were found to be metabolized e.g. galactose, mannose, N-acetyl-

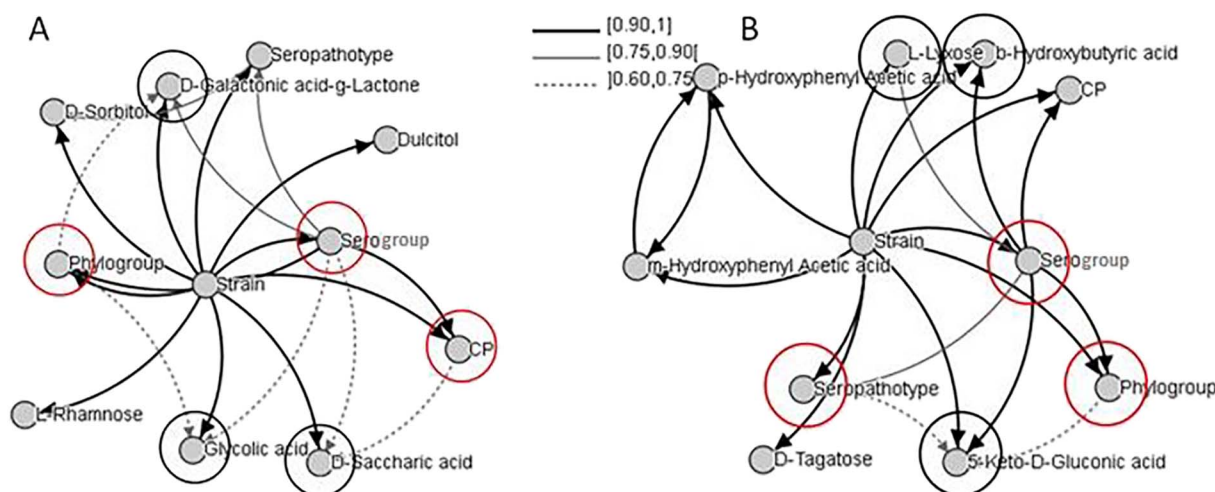


Fig. 6. Adjusted Wallace analysis of the relationships between highly discriminant C-substrates and STEC classifications into serogroups, seropathotypes, and CP (C-utilization patterns) groups. (A) partitioning analysis using the Glycolic acid, D-Galactonic acid- γ -lactone, D-Saccharic acid, D-Sorbitol, Dulcitol, and L-Rhamnose datasets; (B) partitioning analysis using the L-Lyxose, D-Tagatose, 5-Keto-D-Gluconic acid, β -Hydroxybutyric acid, and *m*-Hydroxyphenyl acetic acid datasets. STEC classifications were according Table 1. CP groups are those of Fig. 4. Dashed, grey or black arrows indicate the probability score (see scale on the figure) for the prediction of the state of a character from another one. Most significant relations were circled.

glucosamine (GlcNAc), and *N*-acetyl neuraminic acid (Neu5Ac). Other substrates like maltose and deoxyribose, found to play an important role in O157 (strain EDL933) host colonization (Jones et al., 2008; Schouler et al., 2009), were not used by all strains. In fact, O113 strain 14032 could not use maltose.

3.2. Clustering analysis of C-source utilization patterns

Ratios of C-source utilization patterns (Tables S1 & S2) were converted into a 0/1 matrix which was then used in a clustering analysis by maximum parsimony (MP) allowing a tracking of significant changes between patterns (Fig. 4). The 56 informative substrates identified were used to generate a cladogram. The inferred MP cladogram showed the C-utilization patterns (CP) of the strains to be divided into three large clusters: CP_A ($n = 10$ strains), CP_B ($n = 10$ strains) and CP_C ($n = 17$ strains). The substrates which led to these clusters are indicated in Table 2. Interestingly, each strain had its own specific C-utilization profile.

The CP_A cluster grouped all O157 strains and two O145 strains. This cluster was found matching phylogroup D which is grouping O157 and O145 isolates on the basis of their DNA sequence similarities (Girardeau et al., 2005). Durso et al. (2004) found 28 discriminatory C-substrates between commensal *E. coli* and STEC O157 with D-saccharic acid and D-galactonic acid- γ -lactone being the most informative ones. In our study, these substrates were also found discriminant for O157 of the CP_A cluster. Strains in CP_A were defective in their use. D-galactonic acid- γ -lactone could be a marker of this phylogroup because only MB2653 and PMK5 of O103 were also defective in its use (Fig. S1). Deficiencies in the use of D-saccharic acid were also observed for O26 strains. These latter strains are part of distinct phylogroups, suggesting a convergence in the loss of the metabolic pathways involved in its use. Datasets regarding the use of sorbitol for the CP_A strains were in line with the expected results. All strains but the epidemic sorbitol-fermenting O157:H7 strain 620, were not significantly metabolizing sorbitol (Fig. S1). Similarly, this strain could not use L-sorbose which requires a metabolism by the same pathway (http://www.genome.jp/kegg-bin/show_pathway?map00051+C00247). These datasets also showed, for the first time, that the O157:H7 strain 620 was unable to grow with α -Methyl-D-Galactoside (ratio = 1.5) while most other STEC strains were.

The CP_B cluster grouped profiles of all O26 strains and of two O121 strains. This cluster was found matching phylogroup B1 which is grouping O26 and O121 strains on the basis of their DNA sequences.

However, it did not harbor any of the three O91 strains which are also part of this phylogroup (Girardeau et al., 2005). These datasets support a recent vertical evolution between O26 and O121 strains. Interestingly, CP_B strains were not significantly using D-ribose, lactulose, D-galactopyranosyl-D-arabinose, D-alanine and glycolic acid (glycolate) for their growth. As previously recorded by Hiramatsu et al. (2002), all O26 strains of this study were confirmed as not being able to metabolize rhamnose while most CP_A strains were. Interestingly, another specificity of the O26 strains was their inability at using D-Saccharic acid while the two O121 strains of this cluster were (see outliers on Fig. S1). This deficiency was shared with most of the O157 strains ($n = 6$ over 8), as indicated above. All other STEC strains (except the O104:H4) were able to use D-Saccharic acid for their growth.

The CP_C cluster grouped strains from eight serogroups: O145 ($n = 1$); O111 ($n = 2$); O91 ($n = 3$); O103 ($n = 4$); O113 ($n = 3$); O104 ($n = 1$); O121 ($n = 1$) and O45 ($n = 1$). L-lyxose was found to discriminate strains belonging to the CP_C cluster (Fig. S1). All CP_C strains but 13199 (O91) and 13137 (O113) were able to significantly metabolize L-Lyxose, whereas only one strain from the other clusters metabolized this carbohydrate. Moreover, CP_C strains were all able to use D-Saccharic acid except Zucker-O104 (Fig. S1). These distinct C-source utilization patterns could reflect differences in the survival of STEC among outdoor ecological niches as observed by Franz et al. (2011) for manure. They could also highlight adaptations for the digestive tract or a use of C-sources from other hosts such as plants as previously observed for dulcitol, which is produced by *Cordylanthus*, *Digitalis*, *Leptorhabdos*, and *Melampyrum* (Nolle et al., 2017).

3.3. Correlations between C-source utilization patterns and variations in INT reduction values between highly discriminant C-substrates

To further investigate the use of particular C-substrates as markers of STEC strain groupings, quantitative analyses were performed from the measured area under the curves representing the INT reduction activities. Correlation tests of these datasets between pair of strains were computed (Fig. 5), and confirmed a segregation according to the MP clusters (Fig. 4). These correlation tests were performed on the 56 informative substrates indicated in Table 2. Significant Spearman p -values were found among each cluster (CP_A, CP_B, and CP_C). INT reduction values from the informative substrates were correlated between O157 strains, and between O26 strains (Fig. 5). Boxplot analyses and Wilcoxon Mann Whitney statistical tests were performed to further validate these segregation of INT values according to MP clusters (Fig.

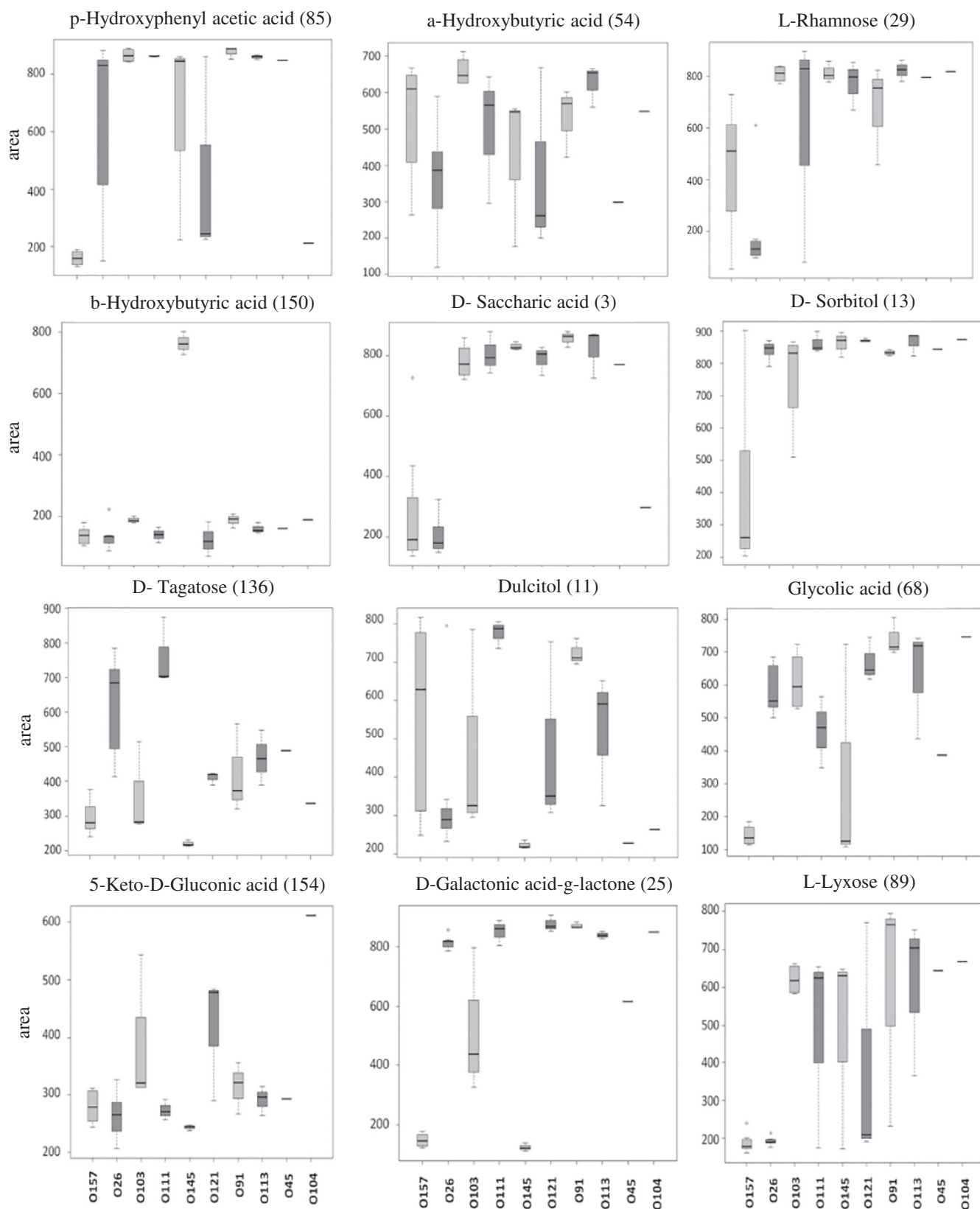


Fig. 7. Boxplot representations of area under the curves computed from INT-reduction activities over a 48 h incubation period per serogroup and for the 12 most discriminant substrates: D-Saccharic acid; Dulcitol; D-Sorbitol; D-Galactonic acid-γ-lactone; L-Rhamnose; α-Hydroxybutyric acid; Glycolic acid; p-Hydroxyphenyl acetic acid; L-Lyxose; D-Tagatose; β-Hydroxybutyric acid; and 5-Keto-D-Gluconic acid. Each box indicates the variations between the first and third quarter of the obtained values, and the median value (thick black line). Outliers are indicated by circles. Computed area under the curves were normalized and transformed prior performing these analyses (see [Materials and methods](#)).

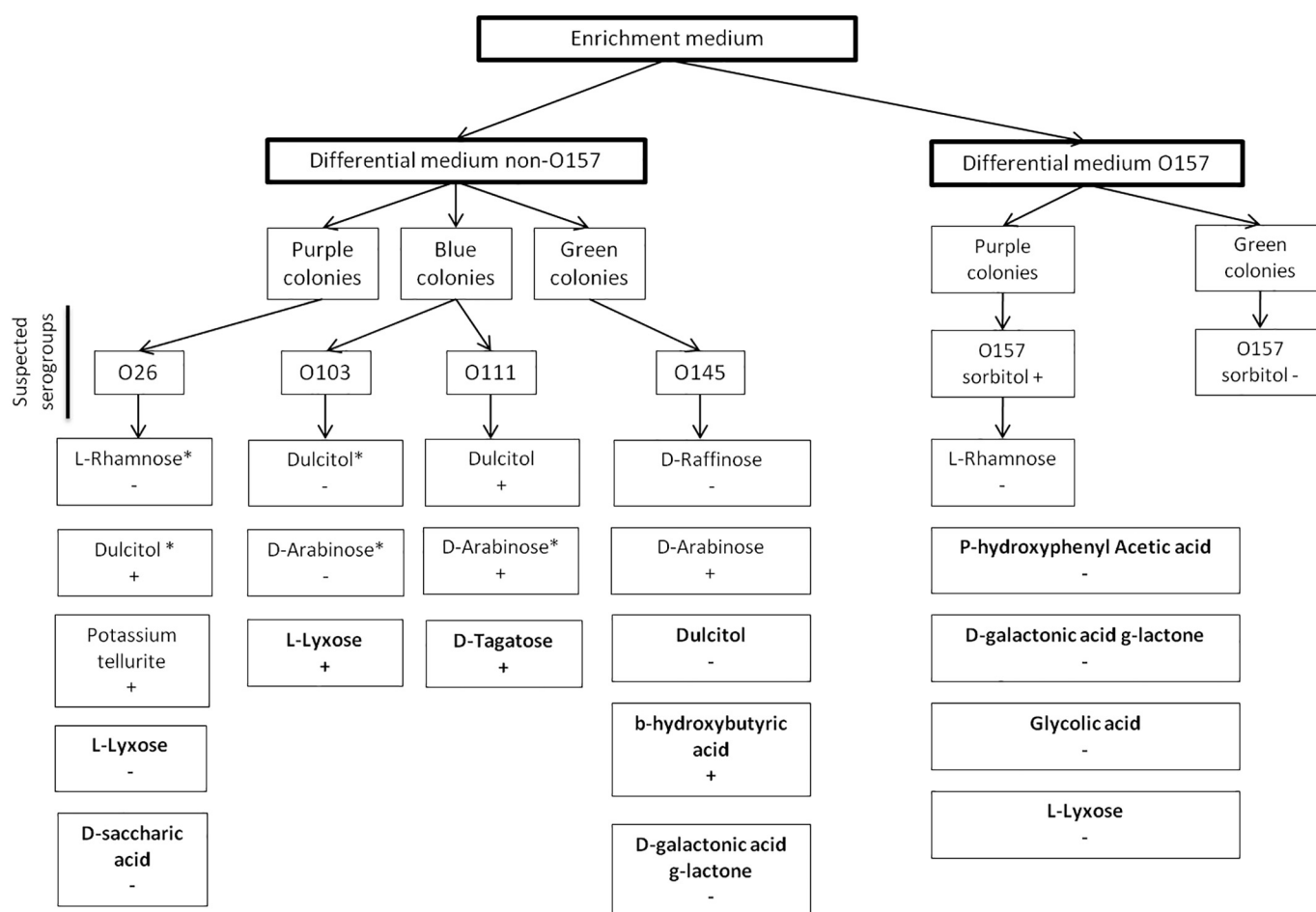


Fig. 8. Identification keys of O26, O45, O91, O103, O104, O111, O113, O121, O145, and O157 STEC strains according to Possé et al. (2008) and Kerangart et al. (2016) completed with discriminatory C-sources identified in this study. Substrates in bold are the ones identified in this work. A star indicates variability in the OM datasets among a same serogroup.

S1). Three substrates were used to illustrate these tests: L-Lyxose, a positive marker for CP_C strains, Sorbitol, a negative marker for CP_A, and D-saccharic acid, a positive marker for CP_C strains but a negative one for O26 strains of the CP_B cluster. A clear differentiation of INT values obtained from L-Lyxose and computed from area under the curves, after 48 h of incubation, were observed, and matched the expected discrimination of CP_C strains from the others. Furthermore, the two CP_C strains (13199 (O91) and 13137 (O113)) not showing a significant use (INT reduction ratios < 2) of L-Lyxose were identified as outliers in the boxplot analysis (Fig. S1). The O145 strain 2513-21 identified as the only user of L-Lyxose among the CP_A cluster was also detected as an outlier (Fig. S1). Similar analyses performed on the datasets for D-Saccharic confirmed the reliability of the clustering approach (Fig. S1). High INT values were obtained for the CP_C cluster, a large variability for the CP_A, and low values for the CP_B cluster while growing STEC strains with D-Saccharic. This analysis identified outliers that were found to match exceptions among the MP cladogram. The Zucker-O104 strain was confirmed a poor grower on D-Saccharic acid, and O121 isolates of CP_B were shown efficient growers while low values were recovered for other strains of this MP cluster. Analysis of INT values related to the use of D-Sorbitol also confirmed the MP clustering analysis (Fig. S1). High INT values were obtained for the CP_C and CP_B strains growing with sorbitol, and a large variability in its use was observed in CP_A one. This latter variability matched the ability or not of O157 strains at using Sorbitol. An outlier was identified among the CP_C cluster and found to match a reduced ability of strain 262.03.18.1 (O103) at using sorbitol, confirming the ratio analysis of Table S1. This analysis further supported the relations observed

between groupings of strains into serogroups and their respective C-metabolic profile.

3.4. Highly discriminant C-substrates matching STEC strain groupings

The reliability of the clustering approach was demonstrated above by independent statistical tests performed on the INT reduction values (area under the curves). This led us to further analyze the MP cladogram and the C-substrates that can differentiate serogroups. Twelve C-substrates among the informative ones were found to have distributions not limited to a few strains, and to give homogeneous results among a same MP cluster or a same serogroup (Table 2). Adjusted Wallace coefficient analyses of the allocation of strains into MP clusters (CP_A, CP_B or CP_C), serogroups, seropathotypes, and phylogroups, and the use or not (binary coding) of the above twelve C-substrates, were performed (Fig. 6). These analyses showed high probabilities for predicting the ability of a serogroup at using glycolic acid, beta-hydroxybutyric acid, L-lyxose and D-galactonic acid-γ-lactone and 5-keto-D-gluconic acid. L-lyxose is a pentose analog of the 6-deoxyhexose L-rhamnose, which can support growth of *E. coli* harboring a mutated rhamnulose kinase favoring a transformation down to glycolate, a salt of glycolic acid (Badia et al., 1991). Glycolic acid is found in several sugar-crops. D-Galactonic acid-γ-lactone is an intermediate in galactose metabolism for some bacteria such as *Pseudomonas* spp. (De Ley and Doudoroff, 1957). β-hydroxybutyric acid occurs widely in nature. Incomplete oxidation of glucose can yield ketogluconates but some *E. coli* can produce 5-keto-dehydrogenases favoring a conversion into D-gluconate (Yum et al., 1999). Interestingly, CP clusters were confirmed by this

approach to be well-predictive of a use of D-saccharic acid by STEC strains (Fig. 6). This latter C-substrate is found in fruits like apples, and vegetables like cabbage and Brussels sprouts. Seropathotypes could not be associated with particular C-substrates using the Wallace analysis. Phylogroups were predictive of the use of D-galactonic acid- γ -lactone and Glycolic acid.

To further explore these trends, boxplot analyses of the area under the curves of the INT reduction values were compared between serogroups for the selected twelve substrates (Fig. 7). Besides, D-sorbitol and L-Rhamnose known to be relatively discriminating, these analyses revealed new specific metabolic capacities per serogroup. The O145 strains were the only ones to use β -hydroxybutyric acid as a C-source. All O111 strains appeared to grow on D-Tagatose while all O157 and O145 could not. Tagatose is a functional sweetener, often found in dairy products. O157 and O45 strains were confirmed not to be able to use D-galactonic acid- γ -lactone and glycolic acid for their growth. No O157 strains grew on p-hydroxyphenyl acetic acid. 5-keto-D-Gluconic acid allowed a specific growth of the O104:H4 strain, and D-Saccharic acid and L-Lyxose did not favor growth of O26 and O157 strains.

3.5. Concluding remarks

STEC identification schemes are based on the use of β -D-galactosidase activity, ability to grow on tellurite and on some C-sources such as sucrose, sorbose, L-Rhamnose, Dulcitol, D-Raffinose, and D-Arabinose (e.g. Possé et al., 2008, and Ratnam et al., 1988). Here, the C-source identification keys of Possé et al. (2008) were confirmed reliable for O145 and O157 but found misleading for O26, O103 and O111 strains. For each of these latter three serogroups, one of the tested strains did not show the expected metabolic capacities (indicated with an asterisk on Fig. 8). This led us to suggest improvements of this scheme by considering the Omnilog PM datasets produced in this study (Fig. 8). The inabilities at using D-Saccharic acid and L-Lyxose were proposed to improve the identification of O26 strains. The positive growth O103 strains with L-Lyxose was suggested to improve its identification (Fig. 8). A positive growth on D-Tagatose was suggested for a differentiation of O111 strains. Dulcitol, β -hydroxybutyric acid and D-Galactonic acid- γ -lactone could be used to improve differentiation of O145 strains. p-Hydroxyphenyl acetic acid, D-Galactonic acid- γ -lactone, Glycolic acid and L-Lyxose were suggested for a differentiation of O157 strains. Similarly, 5-keto-D-Gluconic acid could allow an identification of O104 isolates but this will need to be tested on greater number of strains. Growth tests over larger STEC collections will now need to be performed to validate these proposals.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijfoodmicro.2017.10.018>.

Acknowledgements

This work was partly supported by bioMérieux through a CIFRE grant, and by the French Research Funding Agency (ANR) through the Bacter'eu project (CE04) of the technology for the environment research programme. The French Ministry of Agriculture, CNRS, University Lyon 1 and VetAgro Sup have funded parts of this work through annual grants attributed to the microbial ecology research unit.

We thank R. Marti (BPOE research team of UMR Ecologie Microbienne) for assistance and discussion on the statistical tests used in this work. We thank L. Beutin (National Reference Laboratory for *Escherichia coli* including VTEC, Department of Biological Safety, Federal Institute for Risk Assessment (BfR), Berlin, Germany) for providing *E. coli* strains 12409; 12652; 12805; 12047; 13199; 13694; 13341; 13137 and 14032. We are very grateful to K. Verstraete for having provided reference STEC strains MB 2652 (PH3); MB 2817 (PH88); MB 2653 (PH4) and MB 2659 (PH10). We thank E. Franz for O157 STEC strains m622 and m638. We thank M. Bielaszewska (Institute of Hygiene, National Consulting Laboratory on Hemolytic

Uremic Syndrome, University of Münster, Germany) for *E. coli* O26 human strains 2245/98 and 3073/00, and P. Mariani-Kurkdjian (CNR associé *E. coli*-*Shigella*, Service de Microbiologie, Université Paris-Diderot, Hôpital Robert Debré, Paris, France) for the Zucker-O104:H4 strain. We are also grateful to H. Brugère and D. Bibbal (INSERM UMR1043, INRA USC1360, Université de Toulouse, INP, ENVT, Toulouse, France) for a previous collaborative work that allowed isolating animal strain J43. We finally thank the German NRL team for the *E. coli* field isolates.

References

- Badia, J., Gimenez, R., Balmó, L., Barnes, E., Fessner, W.D., Aguilar, J., 1991. L-lyxose metabolism employs the L-rhamnose pathway in mutant cells of *Escherichia coli* adapted to grow on L-lyxose. *J. Bacteriol.* 173, 5144–5150. <http://dx.doi.org/10.1128/JB.173.16.5144-5150.1991>.
- Bertin, Y., Chaucheyras-Durand, F., Robbe-Masselot, C., Durand, A., de la Foye, A., Harel, J., Cohen, P.S., Conway, T., Forano, E., Martin, C., 2013. Carbohydrate utilization by enterohaemorrhagic *Escherichia coli* O157:H7 in bovine intestinal content. *Environ. Microbiol.* 15, 610–622. <http://dx.doi.org/10.1111/1462-2920.12019>.
- Bettelheim, K. a, Beutin, L., 2003. Rapid laboratory identification and characterization of verocytotoxinigenic (Shiga toxin producing) *Escherichia coli* (VTEC/STEC). *J. Appl. Microbiol.* 95, 205–217.
- Beutin, L., Fach, P., 2014. Detection of Shiga toxin-producing *Escherichia coli* from non-human sources and strain typing. *Microbiol. Spectr.* 2. <http://dx.doi.org/10.1128/microbiolspec.EHEC-0001-2013>.
- Bibbal, D., Loukiadis, E., Kérouredan, M., Ferré, F., Dilasser, F., Peytavin de Garam, C., Cartier, P., Oswald, E., Gay, E., Auvray, F., Brugère, H., 2015. Prevalence of carriage of Shiga toxin-producing *Escherichia coli* serotypes O157:H7, O26:H11, O103:H2, O111:H8, and O145:H28 among slaughtered adult cattle in France. *Appl. Environ. Microbiol.* 81, 1397–1405. <http://dx.doi.org/10.1128/AEM.03315-14>.
- Caprioli, A., Morabito, S., Brugère, H., Oswald, E., 2005. Review article Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission. *Vet. Res.* 36 (2005), 289–311. <http://dx.doi.org/10.1051/vetres>.
- Croen, M.A., Law, R.J., Scholz, R., Keeney, K.M., Włodarska, M., Finlay, B.B., 2013. Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clin. Microbiol. Rev.* 26, 822–880. <http://dx.doi.org/10.1128/CMR.00022-13>.
- De Ley, J., Doudoroff, M., 1957. The metabolism of D-galactose in *Pseudomonas saccharophila*. *J. Biol. Chem.* 227, 745–757.
- Durso, L.M., Smith, D., Hutkins, R.W., 2004. Measurements of fitness and competition in commensal *Escherichia coli* and *E. coli* O157: H7 strains. *Appl. Environ. Microbiol.* 70, 6466–6472. <http://dx.doi.org/10.1128/AEM.70.11.6466>.
- Espié, E., Vaillant, V., Mariani-Kurkdjian, P., Grimont, F., Martin-Schaller, R., De Valk, H., Vernoy-Rozand, C., 2006. *Escherichia coli* O157 outbreak associated with fresh unpasteurized goats' cheese. *Epidemiol. Infect.* 134, 143–146. <http://dx.doi.org/10.1017/S0950268805004887>.
- Farrokh, C., Jordan, K., Auvray, F., Glass, K., Oppegaard, H., Raynaud, S., Thevenot, D., Condron, R., De Reu, K., Govaris, A., Heggum, K., Heyndrickx, M., Hummerjohann, J., Lindsay, D., Miszczyca, S., Moussiegt, S., Verstraete, K., Cerf, O., 2013. Review of Shiga-toxin-producing *Escherichia coli* (STEC) and their significance in dairy production. *Int. J. Food Microbiol.* 162, 190–212. <http://dx.doi.org/10.1016/j.ijfoodmicro.2012.08.008>.
- Fegan, N., Gobius, K.S., 2012. Pathogenic *Escherichia coli* and one health implications. In: Current Topics in Microbiology and Immunology, pp. 49–62. http://dx.doi.org/10.1007/82_2012_261.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution (N. Y.)* 39, 783. <http://dx.doi.org/10.2307/2408678>.
- Feng, P., 2014. Shiga toxin-producing *Escherichia coli* (STEC) in Fresh Produce—a food safety dilemma. *Microbiol. Spectr.* 2 (EHEC-0010-2013). <http://dx.doi.org/10.1128/microbiolspec.EHEC-0010-2013>.
- Franz, E., Hoek, A.H.A.M., Van Bouw, E., Aarts, H.J.M., 2011. Variability of *Escherichia coli* O157 strain survival in manure-amended soil in relation to strain origin, virulence profile, and carbon nutrition profile. *Appl. Environ. Microbiol.* 77 (22), 8088–8096. <http://dx.doi.org/10.1128/AEM.00745-11>. (November 2011).
- Franzin, F.M., Sircilli, M.P., Franzin, F.M., Sircilli, M.P., 2015. Locus of enterocytic effacement: a pathogenicity Island involved in the virulence of Enteropathogenic and Enterohemorrhagic *Escherichia coli* subjected to a complex network of gene regulation. *Biomed. Res. Int.* 2015, 1–10. <http://dx.doi.org/10.1155/2015/534738>.
- Girardeau, J.P., Dalmasso, A., Bertin, Y., Ducrot, C., Bord, S., Livrelli, V., Vernoy-Rozand, C., Martin, C., 2005. Association of virulence genotype with phylogenetic background in comparison to different seropathotypes of Shiga toxin-producing *Escherichia coli* isolates. *J. Clin. Microbiol.* 43, 6098–6107. <http://dx.doi.org/10.1128/JCM.43.12.6098-6107.2005>.
- Grad, Y.H., Lipsitch, M., Feldgarden, M., Arachchi, H.M., Cerqueira, G.C., Fitzgerald, M., Godfrey, P., Haas, B.J., Murphy, C.I., Russ, C., Sykes, S., Walker, B.J., Wortman, J.R., Young, S., Zeng, Q., Abouelleil, A., Bochicchio, J., Chauvin, S., Desmet, T., Gujja, S., McCowan, C., Montmayeur, A., Steelman, S., Frimodt-Møller, J., Petersen, A.M., Struve, C., Krogfelt, K. a, Bingen, E., Weill, F.-X., Lander, E.S., Nussbaum, C., Birren, B.W., Hung, D.T., Hanage, W.P., 2012. Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe, 2011. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3065–3070. <http://dx.doi.org/10.1073/pnas.1214911109>.
- Hayashi, T., 2001. Complete genome sequence of Enterohemorrhagic *Escherichia coli*

- O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res.* 8, 11–22. <http://dx.doi.org/10.1093/dnares/8.1.11>.
- Hiramatsu, R., Matsumoto, M., Miwa, Y., Suzuki, Y., Saito, M., Miyazaki, Y., 2002. Characterization of Shiga toxin-producing *Escherichia coli* O26 strains and establishment of selective isolation media for these strains. *J. Clin. Microbiol.* 40, 922–925.
- Ito, H., Terai, A., Kurazono, H., Takeda, Y., Nishibuchi, M., 1990. Cloning and nucleotide sequencing of Vero toxin 2 variant genes from *Escherichia coli* O91 : H21 isolated from a patient with the hemolytic uremic syndrome. *Microb. Pathog.* 8, 47–60.
- Jones, S.A., Jorgensen, M., Chowdhury, F.Z., Rodgers, R., Hartline, J., Leatham, M.P., Struve, C., Krogfelt, K.A., Cohen, P.S., Conway, T., 2008. Glycogen and maltose utilization by *Escherichia coli* O157:H7 in the mouse intestine. *Infect. Immun.* 76, 2531–2540. <http://dx.doi.org/10.1128/IAI.00096-08>.
- Kaper, J.B., Nataro, J.P., Mobley, H.L., 2004. Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* 2, 123–140. <http://dx.doi.org/10.1038/nrmicro818>.
- Karmali, M.A., 2009. Host and pathogen determinants of verocytotoxin-producing *Escherichia coli*-associated hemolytic uremic syndrome. *Kidney Int. Suppl.* S4–7. <http://dx.doi.org/10.1038/ki.2008.608>.
- Karmali, M.A., Mascarenhas, M., Shen, S., Ziebell, K., Johnson, S., Reid-Smith, R., Isaac-Newton, J., Clark, C., Rahn, K., Kaper, J.B., 2003. Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* serotypes that are linked to epidemic and/or serious disease. *J. Clin. Microbiol.* 41, 4930–4940. <http://dx.doi.org/10.1128/JCM.41.11.4930>.
- Karmali, M. a, Gannon, V., Sargeant, J.M., 2010. Verocytotoxin-producing *Escherichia coli* (VTEC). *Vet. Microbiol.* 140, 360–370. <http://dx.doi.org/10.1016/j.vetmic.2009.04.011>.
- Kerangart, S., Douëllou, T., Delannoy, S., Fach, P., Beutin, L., Sergent-Thévenot, D., Cournoyer, B., Loukiadis, E., 2016. Variable tellurite resistance profiles of clinically-relevant Shiga toxin-producing *Escherichia coli* (STEC) influence their recovery from foodstuffs. *Food Microbiol.* 59, 32–42. <http://dx.doi.org/10.1016/j.fm.2016.05.005>.
- King, L.A., Mailles, A., Mariani-Kurkdjian, P., Vernoy-Rozand, C., Montet, M.P., Grimont, F., Pihier, N., Devalk, H., Perret, F., Bingen, E., Espié, E., Vaillant, V., 2009. Community-wide outbreak of *Escherichia coli* O157:H7 associated with consumption of frozen beef burgers. *Epidemiol. Infect.* 137, 889–896. <http://dx.doi.org/10.1017/S0950268808001490>.
- King, L.A., Nogareda, F., Weill, F.-X., Mariani-Kurkdjian, P., Loukiadis, E., Gault, G., Jourdan-DaSilva, N., Bingen, E., Mace, M., Thevenot, D., Ong, N., Castor, C., Noel, H., Van Cauteren, D., Charron, M., Vaillant, V., Aldabe, B., Goulet, V., Delmas, G., Couturier, E., Le Strat, Y., Combe, C., Delmas, Y., Terrier, F., Vendrely, B., Rolland, P., de Valk, H., 2012. Outbreak of Shiga toxin-producing *Escherichia coli* O104:H4 associated with organic fenugreek sprouts, France, June 2011. *Clin. Infect. Dis.* 54, 1588–1594. <http://dx.doi.org/10.1093/cid/cis255>.
- King, L.A., Loukiadis, E., Mariani-Kurkdjian, P., Haeghebaert, S., Weill, F.-X., Balieri, C., Ganet, S., Gouali, M., Vaillant, V., Pihier, N., Callon, H., Novo, R., Gaillot, O., Thevenot-Sergent, D., Bingen, E., Chaud, P., de Valk, H., 2014. Foodborne transmission of sorbitol-fermenting *Escherichia coli* O157:[H7] via ground beef: an outbreak in northern France, 2011. *Clin. Microbiol. Infect.* 20, O1136–44. <http://dx.doi.org/10.1111/1469-0691.12736>.
- Maddison, W., 2005. *MacClade 4: Analysis of Phylogeny and Character Evolution*. 4. pp. 08a.
- Majowicz, S.E., Diplock, K.J., Leatherdale, S.T., Bredin, C.T., Rebello, S., Hammond, D., Jones-Bitton, A., Dubin, J.A., 2016. Food safety knowledge, attitudes and self-reported practices among Ontario high school students. *Can. J. Public Health* 106, e520–6.
- March, S.B., Ratnam, S., 1986. Sorbitol-MacConkey Medium for detection of *Escherichia coli* O157: H7 Associated with hemorrhagic colitis. *J. Clin. Microbiol.* 23 (5), 869–872 (May).
- Mariani-Kurkdjian, P., Denamur, E., Milon, A., Picard, B., Cave, H., Lambert-Zechovsky, N., Lohr, C., Goulet, P., Sansonetti, P.J., Elion, J., 1993. Identification of a clone of *Escherichia coli* O103:H2 as a potential agent of hemolytic-uremic syndrome in France. *J. Clin. Microbiol.* 31, 296–301.
- Mathusa, E.C., Chen, Y., Enache, E., Hontz, L., 2010. Non-O157 Shiga toxin-producing *Escherichia coli* in foods. *J. Food Prot.* 73, 1721–1736.
- Mora, A., López, C., Dhab, G., López-Beceiro, A.M., Fidalgo, L.E., Díaz, E.A., Martínez-Carrasco, C., Mamani, R., Herrera, A., Blanco, J.E., Blanco, M., Blanco, J., 2012. Serotypes, Phylogroups, Stx subtypes, and intimin types of wildlife-carried, shiga toxin-producing *Escherichia coli* strains with the same characteristics as human-pathogenic isolates. *Appl. Environ. Microbiol.* 78, 2578–2585. <http://dx.doi.org/10.1128/AEM.07520-11>.
- Neto, M., Delannoy, S., Auvray, F., Oswald, E., Fach, P., Loukiadis, E., 2012. Public health significance of *E. coli* O26 isolated from foodstuffs: toward genetic predictors of their virulence. In: Poster presented at the 8th International Symposium on Shiga Toxin (Verocytotoxin) Producing *Escherichia coli* Infections, (Amsterdam, Netherlands).
- Navarro-García, F., 2014. *Escherichia coli* O104:H4 pathogenesis: an enteroregulatory *E. coli*/Shiga toxin-producing *E. coli* explosive cocktail of high virulence. *Microbiol. Spectr.* 2. <http://dx.doi.org/10.1128/microbiolspec.EHEC-0008-2013>.
- Nolle, N., Felsl, A., Heermann, R., Fuchs, T.M., 2017. Genetic characterization of the galactitol utilization pathway of salmonella enterica Serovar *Typhimurium*. *J. Bacteriol.* 199, e00595-16. <http://dx.doi.org/10.1128/JB.00595-16>.
- Ogura, Y., Ooka, T., Iguchi, A., Toh, H., Asadulghani, M., Oshima, K., Kodama, T., Abe, H., Nakayama, K., Kurokawa, K., Tobe, T., Hattori, M., Hayashi, T., 2009. Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 106 (42), 17939–17944. <http://dx.doi.org/10.1073/pnas.0903585106>.
- Perna, N.T., Plunkett, G., Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Pósfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E.J., Davis, N.W., Lim, A., Dimalanta, E.T., Potamousis, K.D., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A., Blattner, F.R., 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 409, 529–533. <http://dx.doi.org/10.1038/35054089>.
- Persad, A.K., LeJeune, J.T., 2014. Animal reservoirs of Shiga toxin-producing *Escherichia coli*. *Microbiol. Spectr.* 2. <http://dx.doi.org/10.1128/microbiolspec.EHEC-0027-2014>.
- Pifer, R., Sperandio, V., 2014. The interplay between the microbiota and enterohemorrhagic *Escherichia coli*. *Microbiol. Spectr.* 2. <http://dx.doi.org/10.1128/microbiolspec.EHEC-0015-2013>.
- Possé, B., De Zutter, L., Heyndrickx, M., Herman, L., 2007. Metabolic and genetic profiling of clinical O157 and non-O157 Shiga-toxin-producing *Escherichia coli*. *Res. Microbiol.* 158, 591–599. <http://dx.doi.org/10.1016/j.resmic.2007.06.001>.
- Possé, B., De Zutter, L., Heyndrickx, M., Herman, L., 2008. Novel differential and confirmation plating media for Shiga toxin-producing *Escherichia coli* serotypes O26, O103, O111, O145 and sorbitol-positive and -negative O157. *FEMS Microbiol. Lett.* 282, 124–131. <http://dx.doi.org/10.1111/j.1574-6968.2008.01121.x>.
- Rasko, D.A., Rosovitz, M.J., Myers, G.S.A., Mongodin, E.F., Fricke, W.F., Gajer, P., Crabtree, J., Sebaihia, M., Thomson, N.R., Chaudhuri, R., Henderson, I.R., Sperandio, V., Ravel, J., 2008. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *J. Bacteriol.* 190, 6881–6893. <http://dx.doi.org/10.1128/JB.00619-08>.
- Ratnam, S., March, S.B., Ahmed, R., Bezanson, G.S., Kasatiya, S., 1988. Characterization of *Escherichia coli* serotype O157:H7. *J. Clin. Microbiol.* 26, 2006–2012.
- Sabarly, V., Bouvet, O., Glodt, J., Clermont, O., Skurnik, B., Chaudhuri, R., de Vienne, D., Denamur, E., Dillmann, C., 2011. The decoupling between genetic structure and metabolic phenotypes in *Escherichia coli* leads to continuous phenotypic diversity. *J. Evol. Biol.* 24, 1559–1571. <http://dx.doi.org/10.1111/j.1420-9101.2011.02287.x>.
- Saxena, T., Kaushik, P., Krishna Mohan, M., 2015. Prevalence of *E. coli* O157:H7 in water sources: an overview on associated diseases, outbreaks and detection methods. *Diagn. Microbiol. Infect. Dis.* 82, 249–264. <http://dx.doi.org/10.1016/j.diagmicrobio.2015.03.015>.
- Schmidt, H., Geitz, C., Tarr, P.I., Frosch, M., Karch, H., 1999. Non-O157:H7 pathogenic *Escherichia coli*: phenotypic and genetic profiling of virulence traits and evidence for clonality. *J. Infect. Dis.* 179, 115–123. <http://dx.doi.org/10.1086/314537>.
- Schouler, C., Taki, A., Chouikha, I., Moulin-Schouleur, M., Gilot, P., 2009. A genomic island of an extraintestinal pathogenic *Escherichia coli* strain enables the metabolism of fructooligosaccharides, which improves intestinal colonization. *J. Bacteriol.* 191, 388–393. <http://dx.doi.org/10.1128/JB.01052-08>.
- Severiano, A., Pinto, F.R., Ramirez, M., Carriço, J.A., 2011. Adjusted Wallace coefficient as a measure of congruence between typing methods. *J. Clin. Microbiol.* 49, 3997–4000. <http://dx.doi.org/10.1128/JCM.00624-11>.
- Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiappello, H., Clermont, O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., El Karoui, M., Frapy, E., Garry, L., Ghigo, J.M., Gilles, A.M., Johnson, J., Le Bouguéne, C., Lescat, M., Mangenot, S., Martinez-Jéhanne, V., Matic, I., Nassif, X., Oztas, S., Petit, M.A., Pichon, C., Rouy, Z., Ruf, C., Saint, Schneider, D., Tourret, J., Vacherie, B., Vallenet, D., Médigue, C., Rocha, E.P.C., Denamur, E., 2009. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet.* 5 (e1000344). <http://dx.doi.org/10.1371/journal.pgen.1000344>.
- Vehkala, M., Shubin, M., Connor, T.R., Thomson, N.R., Corander, J., 2015. Novel R pipeline for analyzing biolog phenotypic microarray data. *PLoS One* 10, e0118392. <http://dx.doi.org/10.1371/journal.pone.0118392>.
- Vernoy-Rozand, C., Montet, M.P., Bertin, Y., Trably, F., Girardeau, J.P., Martin, C., Livrelli, V., Beutin, L., 2004. Serotyping, stx2 subtyping, and characterization of the locus of enterocyte effacement island of Shiga toxin-producing *Escherichia coli* and *E. coli* O157:H7 strains isolated from the environment in France. *Appl. Environ. Microbiol.* 70, 2556–2559. <http://dx.doi.org/10.1128/AEM.70.4.2556-2559.2004>.
- Yum, D.Y., Lee, B.Y., Pan, J.G., 1999. Identification of the yqhE and yafB genes encoding two 2, 5-diketo-D-gluconate reductases in *Escherichia coli*. *Appl. Environ. Microbiol.* 65, 3341–3346.
- Zhang, W.L., Kohler, B., Oswald, E., Beutin, L., Karch, H., Morabito, S., Caprioli, A., Suerbaum, S., Schmidt, H., 2002. Genetic diversity of intimin genes of attaching and effacing *Escherichia coli* strains. *J. Clin. Microbiol.* 40, 4486–4492. <http://dx.doi.org/10.1128/JCM.40.12.4486-4492.2002>.
- Zimmerhackl, L.-B., Rosales, A., Hofer, J., Riedl, M., Jungtraithmayr, T., Mellmann, A., Bielaszewska, M., Karch, H., 2010. Enterohemorrhagic *Escherichia coli* O26:H11-associated hemolytic uremic syndrome: bacteriology and clinical presentation. *Semin. Thromb. Hemost.* 36, 586–593. <http://dx.doi.org/10.1055/s-0030-1262880>.