



OPEN

SUBJECT AREAS:

CLINICAL
MICROBIOLOGY

STATISTICAL METHODS

PATHOGENS

INFRARED SPECTROSCOPY

Received
3 July 2013Accepted
1 November 2013Published
20 November 2013Correspondence and
requests for materials
should be addressed to
L.P. (lpeixe@ff.up.pt)

Diverse high-risk B2 and D *Escherichia coli* clones depicted by Fourier Transform Infrared Spectroscopy

Clara Sousa¹, Ângela Novais¹, Ana Magalhães¹, João Lopes² & Luísa Peixe¹

¹REQUIMTE, Laboratório de Microbiologia, Departamento de Ciências Biológicas, Faculdade de Farmácia, Universidade do Porto, Rua de Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal, ²REQUIMTE, Laboratório de Química Aplicada, Departamento de Ciências Químicas, Faculdade de Farmácia, Universidade do Porto, Rua de Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal.

We aimed to develop a reliable method based on Fourier transform infrared spectroscopy with attenuated total reflectance (FTIR-ATR) to discriminate *Escherichia coli* clones from B2 (n = 9) and D (n = 13) phylogenetic groups. Eighty-eight *E. coli* isolates belonging to phylogenetic groups B2 (n = 39) and D (n = 49), including particularly widespread high risk clones or clonal complexes (HiRCC) ST131, ST69, ST393 and ST405 were studied. Spectra were analysed by unsupervised (hierarchical cluster analysis-HCA) and supervised methods (soft independent modelling of class analogy-SIMCA and partial least square discriminant analysis-PLSDA). B2-ST131 isolates were discriminated from B2 non-ST131 and D phylogroup isolates (ST69, ST393, ST405) by HCA, SIMCA and PLSDA. D-ST69, D-ST393 and D-ST405 isolates were also distinguished from each other and from other STs from phylogroup D by the three methods. We demonstrate that FTIR-ATR coupled with chemometrics is a reliable and alternative method to accurately discriminate particular *E. coli* clones. Its validation towards an application at a routine basis could revolutionize high-throughput bacterial typing.

The global dissemination in different settings of antibiotic resistant high-risk *Escherichia coli* clones or clonal complexes (HiRCC) with high virulence potential constitutes one of the major current challenges in clinical microbiology^{1,2}. Particular *E. coli* clones from phylogenetic groups B2 (ST131) and D (ST69, ST393, ST405) with enhanced ability to colonize, persist and adapt to different hosts are recognized as extraintestinal pathogenic *E. coli* (ExPEC) lineages, which have largely contributed to the dissemination of β -lactam resistance determinants (mainly extended-spectrum β -lactamases and/or carbapenemases) in different countries^{3,4}. Assessing the prevalence and dynamics of ST131 and other HiRCC by quick methods in the clinical setting would have a significant value for clinical, infection control and epidemiological purposes².

Pulsed-field gel electrophoresis (PFGE) and multilocus sequence typing (MLST) have been useful for identification and discrimination of these *E. coli* clones at both local and global levels³⁻⁵, and although other alternative genotyping methods such as multilocus variable number of tandem repeats analysis (MLVA), allele-specific⁶⁻⁹ or real-time PCR¹⁰ or two-locus clonal typing have recently been proposed¹¹, these methods are still time-consuming, laborious and/or expensive. Spectroscopic techniques such as Fourier transform infrared spectroscopy (FTIR) coupled with chemometric tools have demonstrated an interesting potential for the identification and typing of pathogenic and/or antibiotic resistant Gram positive and Gram negative bacteria at different taxonomic levels (species, subspecies, serotype and more recently at the strain level)¹²⁻¹⁸.

The basis of FTIR spectroscopy is the interaction of infrared radiation with a sample, in our specific case with the bacterial isolate, providing a specific fingerprint that reflects the structure and composition of the whole cell¹⁹. In the ATR mode, the infrared beam contact with the bacterial isolate and became attenuated. The magnitude of the attenuation depends on the bacteria in contact with the beam. The main advantages of FTIR-ATR spectroscopy are rapidity and reduced cost. This methodology requires no reagents or only low amounts of consumables, is non-destructive and environmentally friendly²⁰.

In this study, we demonstrate the suitability of FTIR-ATR spectroscopy as a reliable alternative to discriminate diverse *E. coli* clones belonging to phylogenetic groups B2 (n = 9) and D (n = 13), including the particularly widespread ST131, ST69, ST393 and ST405 HiRCC.

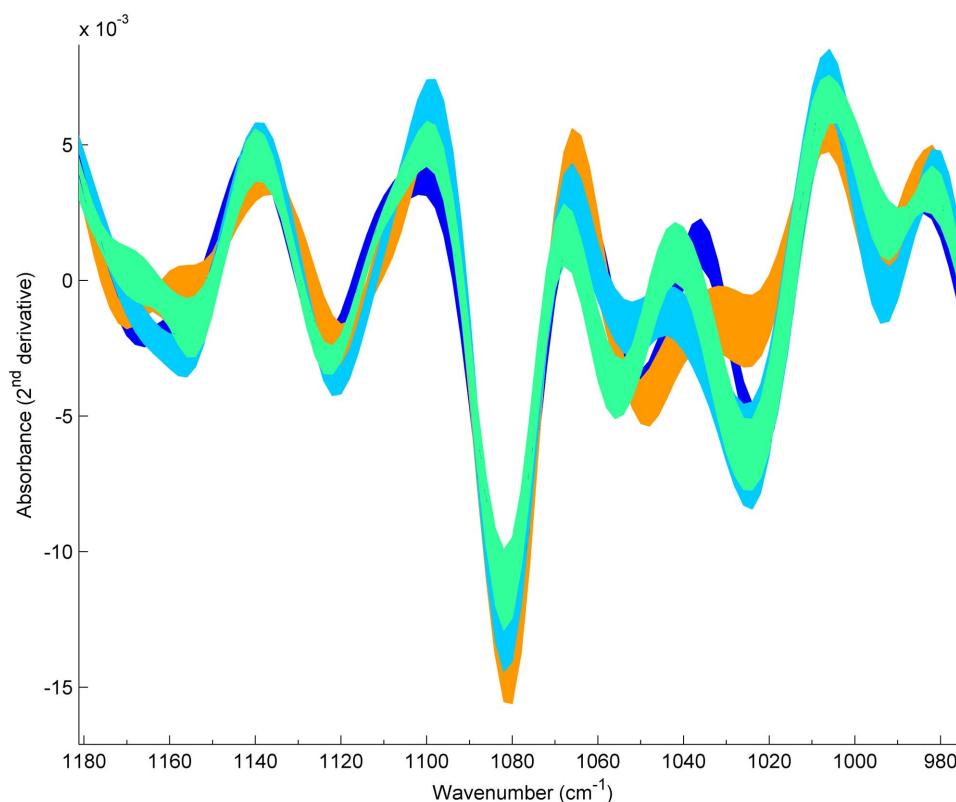


Figure 1 | *Escherichia coli* FTIR-ATR spectra processed with SNV and Savitzky-Golay (9 points filter size, 2nd degree polynomial, 2nd derivative) corresponding to the mean \pm one standard deviations in the region 1180–980 cm^{-1} . Legend: ■ ST131, ■ ST69, ■ ST393 and ■ ST405.

Results

Isolates discrimination work-flow. The rationale of this study was firstly to discriminate the HiRCC B2-ST131 from the phylogenetic group D isolates. After, isolates from phylogroup B2 belonging to diverse STs were included to test FTIR-ATR ability in the discrimination of the ST131 from other B2 phylogroup isolates. Lastly, the discrimination within phylogenetic group D was evaluated. For this, in a first approach, the HiRCC ST69, ST393 and ST405 were discriminated from each other. Subsequently, it was evaluated the ability of FTIR-ATR to discriminate these HiRCC clones from other diverse STs of the phylogroup D.

Spectra overview. FTIR-ATR spectra of all *E. coli* isolates tested revealed high similarity and bands associated with bacterial components such as lipids (3000–2800 cm^{-1}), proteins/amides I and II (1700–1500 cm^{-1}), phospholipids/DNA/RNA (1500–1185 cm^{-1}), polysaccharides (1185–900 cm^{-1}) and the fingerprint region (900–600 cm^{-1}) were observed¹⁹. The main spectral differences were detected in the phospholipids/DNA/RNA and the polysaccharides regions (1500–900 cm^{-1}), which were subsequently chosen for further comparisons. These regions have previously been used in other studies for discrimination at different taxonomic levels (including clones) in other bacterial species^{16,18,21}.

Discrimination of B2-ST131 isolates. B2-ST131 isolates were clearly discriminated from those belonging to phylogroup D by both HCA and PLSDA chemometric methods. The main spectral differences were observed in the regions of 1165–1155 cm^{-1} , 1115 cm^{-1} and 1050–1030 cm^{-1} (attributed to aromatic vibrations, RNA ribose C-O stretching and diverse carbohydrates vibrations, respectively) (Figure 1)²². The dendrogram generated by HCA grouped all B2-ST131 isolates in one branch (including the single locus variant (SLV) ST1035) and all D *E. coli* clones in another branch containing ST69, ST393 and ST405 isolates (Figure 2).

Consistently, the score plot obtained by PLSDA revealed the discrimination of B2-ST131 isolates by the first latent variable (LV1), which encompasses 24.8% of the total spectral variability (Figure 3). Both methods HCA and PLSDA, presented 100% sensitivity and 100% specificity in the discrimination of B2-ST131 meaning that all isolates of this ST were predicted as ST131 and all D phylogroup isolates were predicted as non B2-ST131 isolates. For sensitivity and specificity calculations the SLV isolates were excluded.

Discrimination within phylogroup B2 isolates. The comparison of FTIR-ATR spectra corresponding to ST131 isolates with those obtained from other B2 *E. coli* clones revealed that ST131 isolates were clearly discriminated from B2 non-ST131 isolates. The dendrogram obtained by HCA evidenced two clusters, one including all ST131 isolates and the SLV ST1035 and the other containing all non-ST131 isolates (Figure 4). A dendrogram generated only with ST131 isolates in the same conditions grouped isolates in two main clusters (data not shown) non-homologous to those observed by genotypic methods²³. In addition, we performed a SIMCA model with ST131 isolates in order to test if the non-ST131 isolates were correctly predicted as not belonging to the modelled ST131 class. The correct discrimination of all the available samples (test samples and the non-ST131 isolates) was achieved with a three component model (Figure 5). All ST131 test samples appeared below the confidence limit confirming their assignment to the B2-ST131 class and B2 non-ST131 isolates appeared above the confidence limit meaning that they do not belong to that class. Both methods, HCA and SIMCA, presented 100% of sensitivity and specificity in the discrimination of B2-ST131 from B2 non-ST131.

Discrimination within phylogroup D isolates. The comparison of FTIR-ATR spectra of ST69, ST393 and ST405 isolates showed remarkable differences between ST69 and ST405 at 1025 cm^{-1}

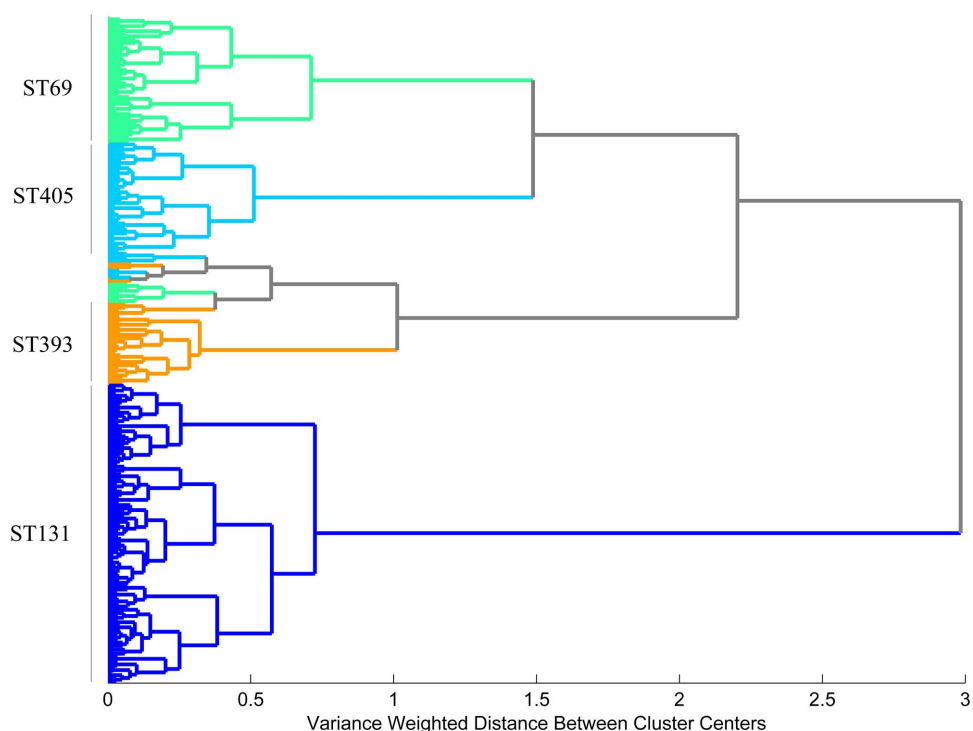


Figure 2 | Dendrogram obtained from the 1500–900 cm^{-1} spectral region using the Ward's algorithm and 9 principal components (PCs) distance for isolates of the B2 (ST131) and D (ST69, ST393 and ST405) phylogenetic groups.

(associated with the S=O stretching of organic sulfoxides) and between ST393 and ST405 at 985 cm^{-1} (corresponding to the asymmetric stretching of $(\text{CH}_3)_3\text{N}^+$) (Figure 1)²². The dendrogram generated by HCA with these isolates revealed that they were

grouped in three clusters, containing respectively most ST69 ($n = 11/13$), all ST393 ($n = 13$) and most ST405 ($n = 9/11$) isolates (Figure 2). In this analysis, we obtained good sensitivity and specificity values for ST69 (85% vs 100%), ST393 (100% vs 83%)

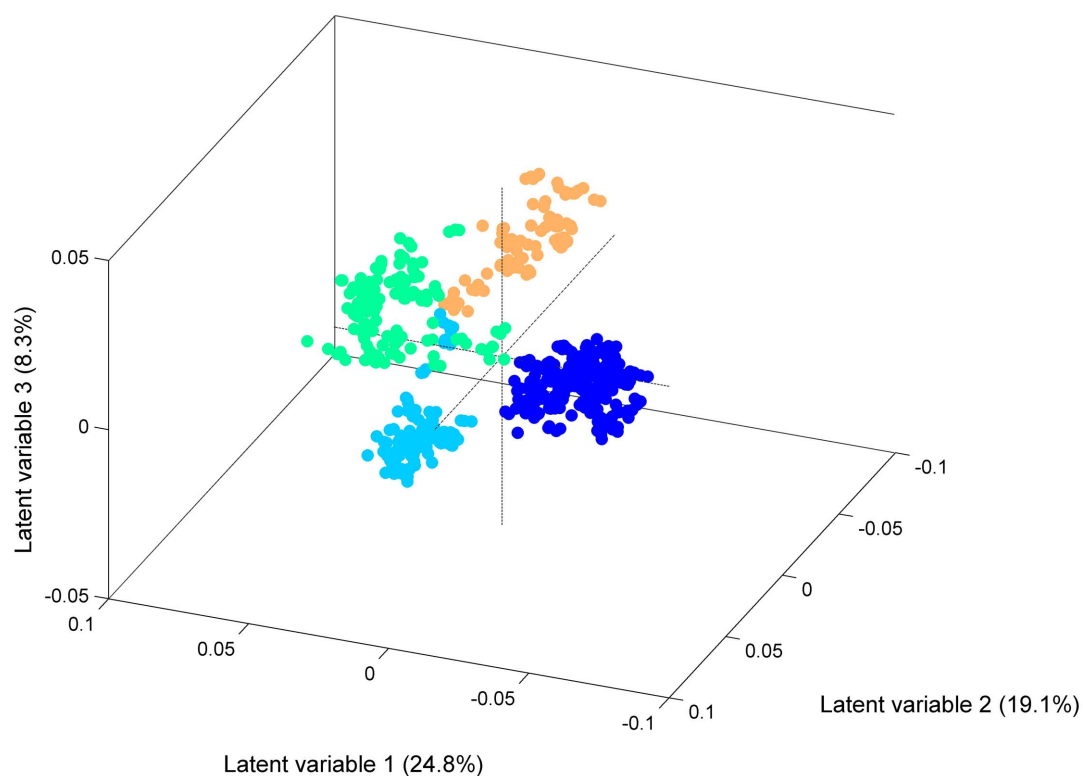


Figure 3 | Score plot corresponding to the three first LVs of the PLS regression model with isolates belonging to phylogenetic groups B2 and D
Legend: ● B2-ST131, ● D-ST69 ● D-ST393 and ● D-ST405.

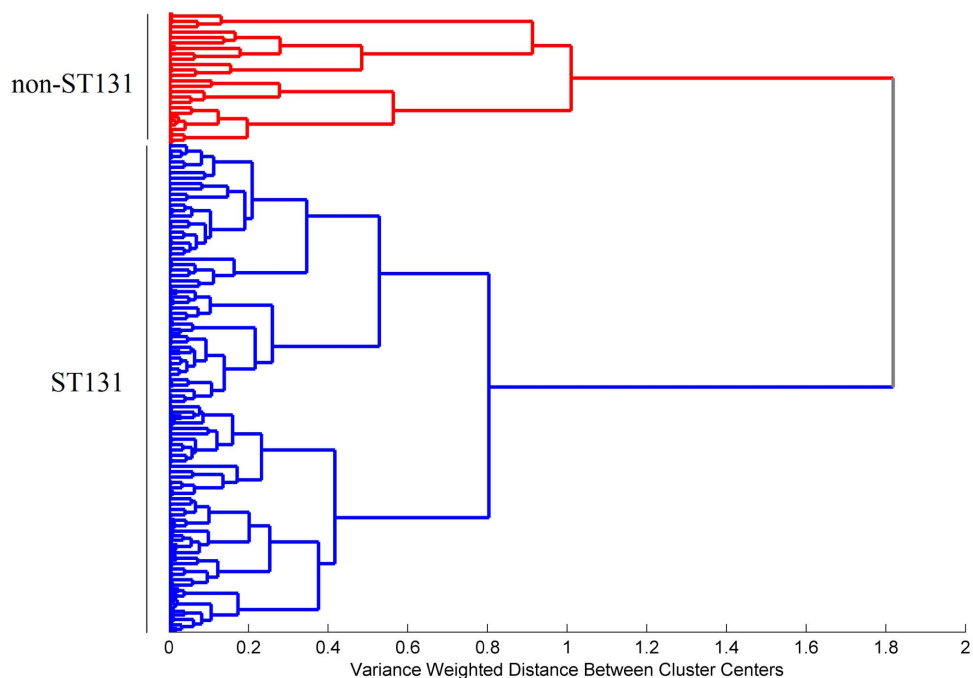


Figure 4 | Dendrogram obtained from the 1500–900 cm^{-1} spectral region using the Ward's algorithm, the Mahalanobis distance and 10 principal components (PCs) for the isolates of the B2 phylogenetic group.

and ST405 (80% vs 100%). However, these clones could be perfectly discriminated by PLSDA (100% of sensitivity and 100% specificity). In fact, three clusters are evidenced in the score map, each one containing isolates from a given ST (ST69, ST393 and ST405), including the respective SLVs (Figure 6).

In addition, we tested the ability of FTIR-ATR to differentiate isolates belonging to ST69, ST393 and ST405 from isolates belonging to other diverse STs of the phylogroup D (hereby designated as other STD). The dendrogram generated by HCA showed that isolates grouped in four different clusters, one of them including the other

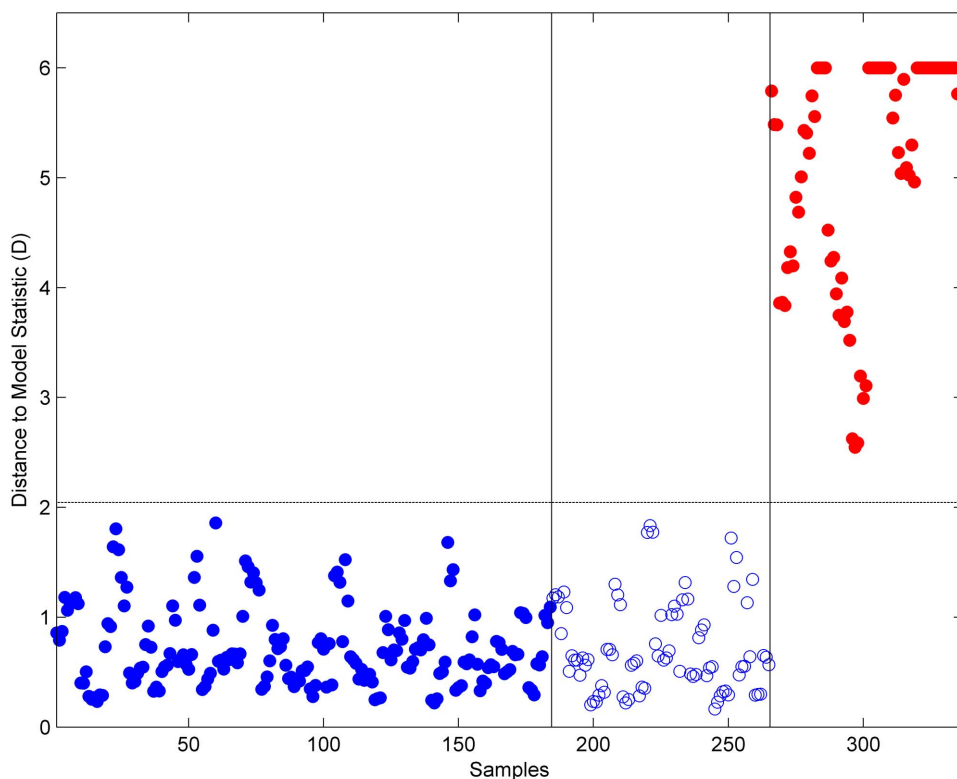


Figure 5 | Distance to model statistics obtained by projecting ST131 and B2 non-ST131 samples on a SIMCA model calibrated with ST131 samples (note that values above 6 are truncated for better visualization). Legend: ● B2-ST131 and ● B2 non-ST131 (Unfilled circles indicated the samples used to test the model).

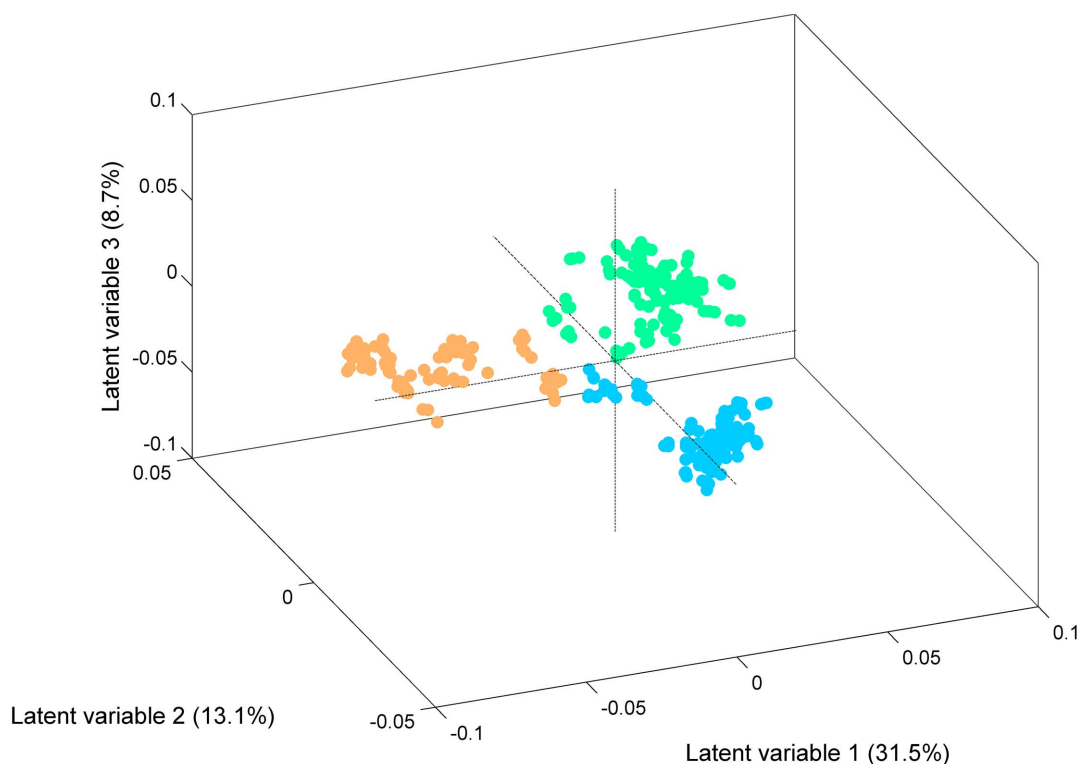


Figure 6 | Score plot corresponding to the three first LVs of the PLSDA regression model with isolates belonging to phylogenetic group D. Legend: ● ST69 ● ST393 and ● ST405.

STD isolates ($n = 14$) (Figure 7). HCA's sensitivity and specificity for the discrimination of ST69, ST393 and ST405 from the other STD isolates was 85% and 100%, respectively. Despite the diversity of STs included in the other STD group, they probably clustered together due to a higher similarity within them than with ST69, ST393 and ST405 isolates. Nothing can be inferred about the relative similarity among the STs of the other STD once just few isolates of each ST was considered in the analysis.

In order to test if all other STD isolates were correctly predicted as not belonging to ST69, ST393 and ST405, three independent SIMCA models were constructed for these three clonal groups. In each

individual ST model, isolates from the remaining STs and from other STD were then projected into the model. All the test samples from ST69, ST393 and ST405 clonal groups were well assigned, whereas other STD isolates were in all cases predicted as not belonging to those groups (Figure 8) meaning that the SIMCA model had 100% of sensitivity and 100% specificity.

Discussion

In this study, we demonstrate that FTIR-ATR spectroscopy coupled with chemometric tools is an alternative and reliable method to accurately discriminate particular *E. coli* clones belonging to B2

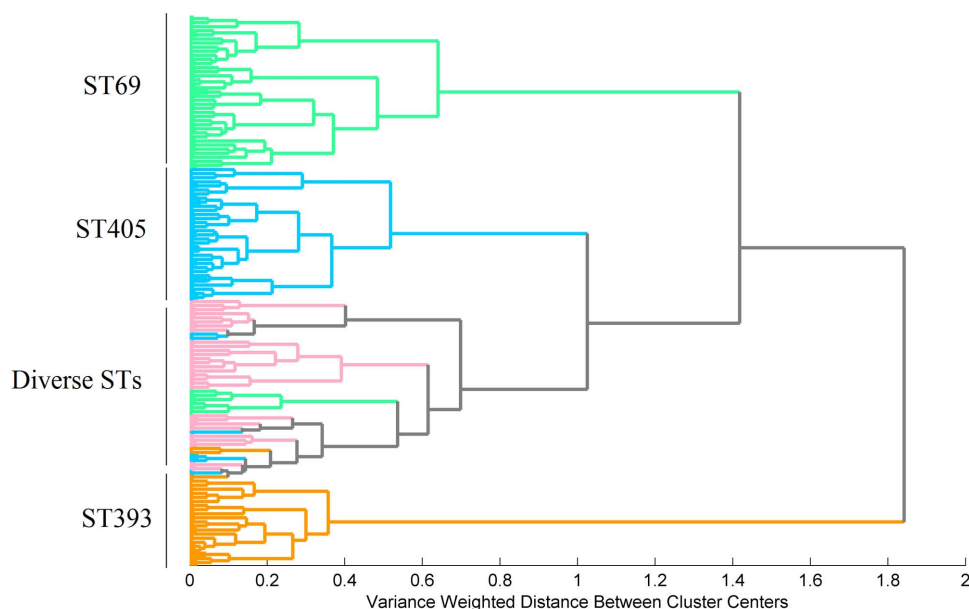


Figure 7 | Dendrogram obtained from the 1500–900 cm^{-1} spectral region using the Ward's algorithm and 17 PCs for the isolates of the phylogroup D.

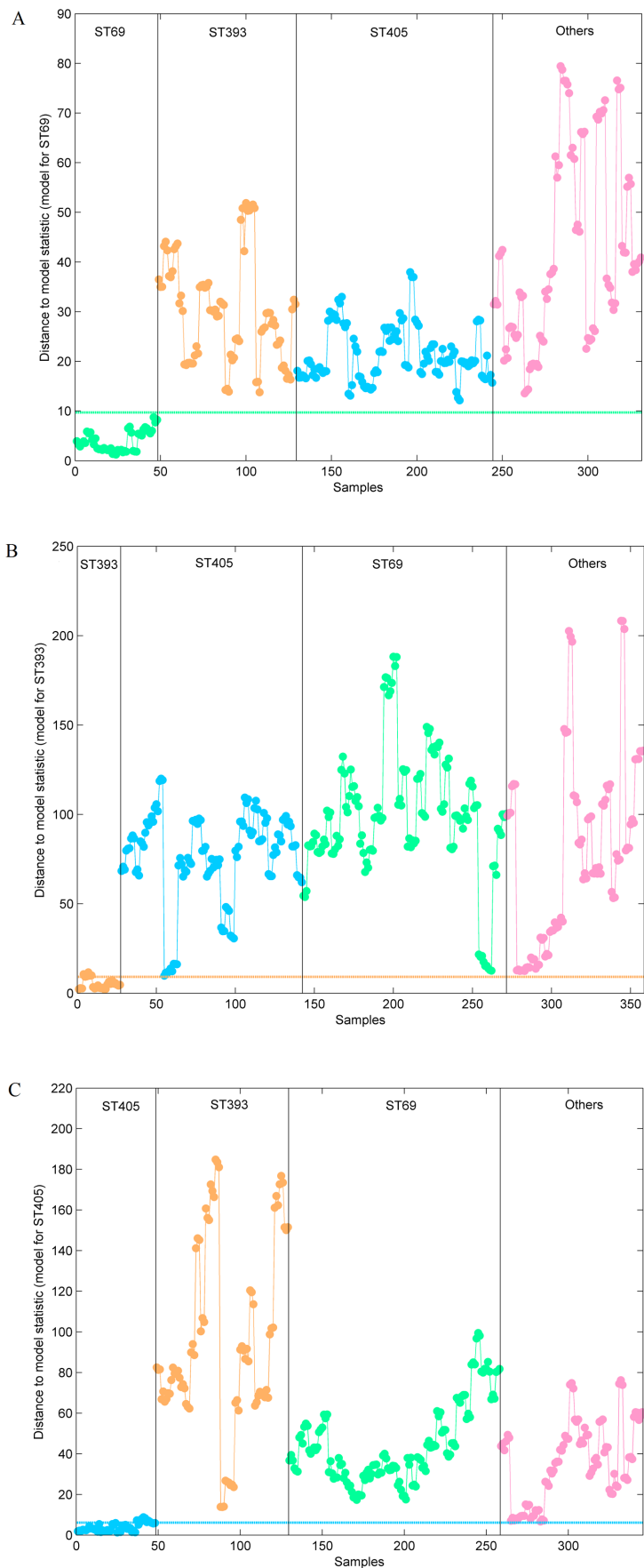


Figure 8 | Distance to model statistics obtained for a SIMCA model for the three STs of phylogenetic group D: A) Model for ST69; B) model for ST393 and C) model for ST405. Legend: ● D-ST69 ● D-ST393 and ● D-ST405 and ● Diverse STs.



Table 1 | Epidemiological details of the *E. coli* isolates used in this work. (H = Hospitalized patients; C = Community patients; F = Healthy volunteers; A = animals; E = Environment; S = Ready-to-eat salads)

Phylogroup	ST/ <i>fumC</i>	Continent (N° countries)	N° of Isolates	Date	Origin ^a (N° Isolates)	Ref	
B2	131	Europe (7)	26	1991–2010	H (19), C (1), F (4), E (2)	[23]	
	131	North America (1)	3	2008	A (3)	[23]	
	131	Asia (1)	2	2006–2007	C (2)	[23]	
	1035	Europe (1)	1	2005	H (1)	[23]	
	95	South America (1)	1	2011	H (1)	*	
	12	Europe (1)	1	2007	H (1)	*	
	355-like	Europe (1)	1	2007	H (1)	*	
	126-like	Europe (1)	1	2007	A (1)	[25]	
	799-like	Europe (1)	1	2006/07	A (1)	[25]	
	<i>fumC11</i>	Europe (1)	1	2010	H (1)	*	
	<i>fumC103</i>	Europe (1)	1	2006	H (1)	*	
	D	69	Europe (3)	7	2002–2010	H (3), S (2), C (1), A (1)	[24]
		69	North America (1)	5	1999–2000	H (5)	[24]
69		South America (1)	1	-	H (1)	[24]	
393		Europe (3)	5	2002–2006	F (3), C (1), H (1)	[24]	
393		North America (1)	3	1980–1999	H (3)	[24]	
393		Asia (1)	2	2006–7	C (2)	[24]	
2321		Europe (1)	1	2008	H (1)	[24]	
405		Europe (4)	9	2000–2008	H (7), C (2)	[24]	
405		Asia (1)	1	2004	H (1)	[24]	
964		Europe (1)	1	2003	H (1)	[24]	
117		Europe (1)	4	2006–2010	H (4)	*	
New ST ^a		Europe (1)	1	2010	H (1)	*	
648		Europe (1)	1	2006	H (1)	*	
1011		Europe (1)	2	2010	H (2)	*	
1325		Europe (1)	1	2010	H (1)	*	
3177		Europe (1)	1	2007	H (1)	*	
<i>fumC88</i>		Europe (1)	1	2010	H (1)	*	
<i>fumC31</i>	Europe (1)	3	2010	H (3)	*		

^aDouble locus variant (DLV) of ST117;
*unpublished data.

and D phylogenetic groups. The method proposed here consists on a reproducible framework where we discriminate sequentially: i) B2-ST131 from D-ST69, D-ST393 and D-ST405 *E. coli* clones; ii) B2-ST131 from other B2 *E. coli* clones; iii) ST69, ST393 and ST405; and iv) ST69, ST393 and ST405 from other D *E. coli* clones.

Strategies to shorten the time for the detection of these multidrug resistant and virulent clones and/or with potential application at a large-scale basis are being increasingly pursued^{6,11} since their application at a routine basis would have relevant clinical, infection control and epidemiological implications. Despite the increasing number of applications at different taxonomic levels, FTIR has very rarely been tested for clonal differentiation^{16,17,21}. The study by AlRabiah et al. demonstrates the ability of FTIR to discriminate a few *E. coli* isolates involved in urinary tract infections including members of the ST131 clone. In our study, we included a larger and diverse sample of previously characterized *E. coli* isolates from different clones, demonstrating that FTIR might constitute a new and promising field in high-throughput bacterial typing. The results obtain herein proved FTIR precision (high consistency between biological and instrumental replicates) and accuracy (correct clonal group prediction).

FTIR provides a whole organism fingerprint¹⁹ that appears to be related with its phenotypic and genotypic features, since a good correlation was found between the assignments obtained by comparison of FTIR spectra and the STs determined by MLST, as observed previously in *Acinetobacter baumannii*¹⁶. FTIR seems to have a lower discriminatory power than MLST since SLVs of a given ST were not recognized, suggesting similarity of genotypic or phenotypic characters, as previously observed. However, isolates' clusterization did not correlate with that obtained by PFGE or with similarity of antibiotic resistance or virulence gene profiles^{23,24}, suggesting that FTIR is possibly depicting more stable features. Its reliability for clonal discrimination will be further tested in a higher

number of *E. coli* isolates and the possibility to extend the discriminatory power to other *E. coli* clones and eventually to other *E. coli* phylogenetic groups will be further explored.

FTIR spectrometers are available in many academic departments, laboratory research units or industries for a variety of purposes in chemistry and biochemistry such as characterization and quantification of chemical compounds or drugs, real time process monitoring or identification of potential bio threats or toxics²⁰. The potential of this equipment for other goals may have been neglected over the years. We believe that if spectral acquiring conditions and the same equipment are assured, this method could be suitable for routine implementation in other laboratories enabling quickly and at a low cost the detection of high-risk *E. coli* clones, which would positively influence individual patient management decisions, infection control measures and monitorization of epidemiological trends. Finally, FTIR could be proposed as a reliable alternative to discriminate particular *E. coli* clones from B2 and D phylogenetic groups revolutionizing clinical bacteriology routines and high-throughput bacterial typing.

Methods

Bacterial strains. A set of eighty-eight *E. coli* isolates belonging to 22 clones from B2 (31 ST131, 1 ST1035, 1 ST12, 1 ST95, 1 ST126-like, 1 ST355-like, 1 ST799-like, 1 *fumC12*, 1 *fumC103*) and D (13 ST69, 10 ST393, 10 ST405, 1 ST2321, 1 ST964, 4 ST117, 1 newST, 1 ST648, 2 ST1011, 1 ST1325, 1 ST3177, 1 *fumC88*, 3 *fumC31*) phylogenetic groups were studied. They represent a diversity of previously characterized isolates identified in multiple countries, origins and periods (1980–2010), comprising diverse PFGE-types and variants (sharing identical virulence and/or antibiotic resistance profiles) from each clonal group^{23,24}. Details about the bacterial isolates included in this study are summarized in Table 1.

FTIR spectra acquisition. Spectra were acquired using a PerkinElmer Spectrum BX FTIR System spectrophotometer in the ATR mode with a PIKE Technologies Gladi ATR accessory from 4000–400 cm⁻¹ and a resolution of 4 cm⁻¹ and 32 scan



co-additions. Isolates were grown on Mueller Hinton agar at 37°C for 18 h and colonies were directly applied in the ATR crystal and dried in a thin film. For each isolate, 9 spectra were acquired corresponding to three biological replicates (obtained from the same agar plate) and three instrumental replicates (obtained in three independent days).

Spectral modeling. FTIR-ATR spectra were processed with standard normal variate (SNV)²⁶ followed by the application of a Savitzky-Golay filter (9 smoothing points, 2nd order polynomial and second derivative)²⁷, mean-centred and analysed by unsupervised and supervised chemometric methods. All spectra (nine replicates for each isolate) was considered in the analysis and represented in the figures. The chemometric analysis was performed in Matlab version 6.5 release 13 (MathWorks, Natick, MA) and the PLS Toolbox version 3.5 for Matlab (Eigenvector Research, Manson, WA).

The unsupervised method employed was the hierarchical cluster analysis (HCA)²⁶ using the Ward's algorithm to evaluate spectral similarity. Dendrograms produced by HCA were obtained after a principal component analysis (PCA)²⁶, which ensured the robustness of the results. The supervised methods used were partial least square discriminant analysis (PLSDA)^{28,29} and soft independent modelling of class analogy (SIMCA)³⁰. The PLSDA model is based on the PLS regression method²⁸, and requires a previous knowledge about all the samples used. The model was calibrated considering all samples and the leave-one-sample-out cross-validation procedure in order to prevent overfitting^{30,31}. The SIMCA model is based on the development of multiple PCA models, each one considering data for a particular class, and samples to be classified are then projected onto these models. In each model, 70% of randomly selected isolates of each ST were used for calibration (calibration samples) and 30% for testing (test samples). In our case, samples' class assignment³¹ was performed with the Euclidean distance. This model shows appropriateness when the objective is to classify samples within a defined set of classes and also to identify samples not belonging to any class.

- Woodford, N., Turton, J. F. & Livermore, D. M. Multiresistant Gram-negative bacteria: the role of high-risk clones in the dissemination of antibiotic resistance. *FEMS Microbiol. Rev.* **35**, 736–755 (2011).
- Baquero, F. & Coque, T. M. Multilevel population genetics in antibiotic resistance. *FEMS Microbiol. Rev.* **35**, 705–706 (2011).
- Coque, T. M. *et al.* Dissemination of clonally related *Escherichia coli* strains expressing extended-spectrum beta-lactamase CTX-M-15. *Emerg. Infect. Dis.* **14**, 195–200 (2008).
- Blanco, J. *et al.* National survey of *Escherichia coli* causing extraintestinal infections reveals the spread of drug-resistant clonal groups O25b:H4-B2-ST131, O15:H1-D-ST393 and CGA-D-ST69 with high virulence gene content in Spain. *J. Antimicrob. Chemother.* **66**, 2011–2021 (2011).
- Johnson, J. R. *et al.* Comparison of *Escherichia coli* ST131 pulsotypes, by epidemiologic traits, 1967–2009. *Emerg. Infect. Dis.* **18**, 598–607 (2012).
- Clermont, O. *et al.* Rapid detection of the O25b-ST131 clone of *Escherichia coli* encompassing the CTX-M-15-producing strains. *J. Antimicrob. Chemother.* **64**, 274–277 (2009).
- Johnson, J. R., Owens, K., Sabate, M. & Prats, G. Rapid and specific detection of the O15:K52:H1 clonal group of *Escherichia coli* by gene-specific PCR. *J. Clin. Microbiol.* **42**, 3841–3843 (2004).
- Johnson, J. R., Owens, K., Manges, A. R. & Riley, L. W. Rapid and specific detection of *Escherichia coli* clonal group A by gene-specific PCR. *J. Clin. Microbiol.* **42**, 2618–2622 (2004).
- Naseer, U. *et al.* Multi-Locus Variable Number of Tandem Repeat Analysis for Rapid and Accurate Typing of Virulent Multidrug Resistant *Escherichia coli* Clones. *Plos one* **7**, e41232 (2012).
- Dhanji, H. *et al.* Real-time PCR for detection of the O25b-ST131 clone of *Escherichia coli* and its CTX-M-15-like extended-spectrum beta-lactamases. *Int. J. Antimicrob. Agents* **36**, 355–358 (2010).
- Weissman, S. J. *et al.* High-resolution two-locus clonal typing of extraintestinal pathogenic *Escherichia coli*. *Appl. Environ. Microbiol.* **78**, 1353–1360 (2012).
- Winder, C. L., Gordon, S. V., Dale, J., Hewinson, R. G. & Goodacre, R. Metabolic fingerprints of *Mycobacterium bovis* cluster with molecular type: implications for genotype-phenotype links. *Microbiology* **152**, 2757–2765 (2006).
- Kuhm, A. E., Suter, D., Felleisen, R. & Rau, J. Identification of *Yersinia enterocolitica* at the species and subspecies levels by Fourier transform infrared spectroscopy. *Appl. Environ. Microbiol.* **75**, 5809–5813 (2009).
- Preisner, O., Guiomar, R., Machado, J., Menezes, J. C. & Lopes, J. A. Application of Fourier transform infrared spectroscopy and chemometrics for differentiation of *Salmonella enterica* serovar Enteritidis phage types. *Appl. Environ. Microbiol.* **76**, 3538–3544 (2010).
- Davis, R. & Mauer, L. J. Subtyping of *Listeria monocytogenes* at the haplotype level by Fourier transform infrared (FT-IR) spectroscopy and multivariate statistical analysis. *Int. J. Food Microbiol.* **150**, 140–149 (2011).
- Sousa, C., Grosso, F., Meirinhos-Soares, L., Peixe, L. & Lopes, J. A. Identification of carbapenem-resistant *Acinetobacter baumannii* clones using infrared spectroscopy. *J. Biophotonics*; DOI 10.1002/jbio.201200075 (2012).
- AlRabiah, H., Correa, E., Upton, M. & Goodacre, R. High-throughput phenotyping of uropathogenic *E. coli* isolates with Fourier transform infrared spectroscopy. *Analyst* **138**, 1363–1369 (2013).
- Vaz, M. *et al.* Serotype discrimination of encapsulated *Streptococcus pneumoniae* strains by Fourier-transform infrared spectroscopy and chemometrics. *J. Microbiol. Meth.* **93**, 102–107 (2013).
- Naumann, D., Helm, D. & Labischinski, H. Microbiological characterizations by FT-IR spectroscopy. *Nature* **351**, 81–82 (1991).
- Blum, M. & Harald, J. Historical perspective and modern applications of Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy (ATR-FTIR). *Drug Test Anal.* **4**, 298–302 (2012).
- Coutinho, C. P., Sá-Correia, I. & Lopes, J. A. Use of Fourier transform infrared spectroscopy and chemometrics to discriminate clinical isolates of bacteria of the *Burkholderia cepacia* complex from different species and ribopatterns. *Anal. Bioanal. Chem.* **394**, 2161–2171 (2009).
- Stuart, B. Biological Applications. In: *Infrared Spectroscopy Fundamentals and Applications*. John Wiley & Sons. New Jersey. (2004).
- Novais, Á. *et al.* Characterization of Globally Spread *Escherichia coli* ST131 Isolates (1991 to 2010). *Antimicrob. Agents Chemother.* **56**, 3973–3976 (2012).
- Novais, Á. *et al.* Diversity and biofilm-production ability among isolates of *Escherichia coli* phylogroup D belonging to ST69, ST393 and ST405 clonal groups. *BMC Microbiology*, “in press” (2013).
- Rodrigues, C., Machado, E., Peixe, L. & Novais, Á. IncI1/ST3 and IncN/ST1 plasmids drive the spread of bla_{TEM}-52 and bla_{CTX-M}-1/-32 in diverse *Escherichia coli* clones from different piggeries. *J. Antimicrob. Chemother.* DOI: 10.1093/jac/dkt187 (2013).
- Næs, T., Isaksson, T., Fearn, T. & Davis, T. *A user-friendly guide to multivariate calibration and classification*. NIR Publications. Chichester. (2002).
- Savitzky, A. & Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **36**, 1627–1639 (1964).
- Geladi, P. & Kowalsky, B. R. Partial least-squares regression: a tutorial. *Anal. Chim. Acta* **185**, 1–17 (1986).
- Alsberg, B. K., Kell, D. B. & Goodacre, R. Variable selection in discriminant partial least-squares analysis. *Anal. Chem.* **70**, 4126–4133 (1998).
- Wold, S. Pattern recognition by means of disjoint principal components analysis. *Pattern Recogn.* **8**, 127–139 (1976).
- De Maesschalck, R., Candelari, A., Massart, D. L. & Heuerding, S. Decision criteria for soft independent modeling of class analogy applied to near infrared data. *Chemometr. Intell. Lab.* **47**, 65–77 (1999).

Acknowledgments

This study was supported by FCT grants (PEst-C/EQB/LA0006/2011, POCI/AMB/61814/2004-FSE/FEDER, EXPL/DTP-EPI/0196/2012, FCOMP-01-0124-FEDER-027745 and PTDC/AAC-AMB/103386/2008). Clara Sousa was supported by a post-doctoral grant from Fundação para a Ciência e a Tecnologia (SFRH/BPD/70548/2010). Ângela Novais was supported by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme (PIEF-GA-2009-255512) and an ESCMID Research Grant 2012. The authors declare no conflicts of interest.

Author contributions

C.S., Á.N. and L.P. participated in the conception and design of the study. C.S., Á.N. and A.M. contributed to data acquisition. C.S., Á.N. and J.L. performed data analysis. C.S., Á.N., J.L. and L.P. contributed to the preparation of the manuscript. All the authors read and approved the final manuscript.

Additional information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Sousa, C., Novais, Á., Magalhães, A., Lopes, J. & Peixe, L. Diverse high-risk B2 and D *Escherichia coli* clones depicted by Fourier Transform Infrared Spectroscopy. *Sci. Rep.* **3**, 3278; DOI:10.1038/srep03278 (2013).



This work is licensed under a Creative Commons Attribution 3.0 Unported license. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0>