

Escherichia coli isolates from
extraintestinal organs of livestock animals
harbour diverse virulence genes and
belong to multiple genetic lineages

Article

Accepted Version

Wu, G., Erlicht, R., Mafura, M., Stokes, M., Smith, N.,
Pritchard, G. C. and Woodward, M. J. (2012) *Escherichia coli*
isolates from extraintestinal organs of livestock animals
harbour diverse virulence genes and belong to multiple genetic
lineages. *Veterinary Microbiology*, 160 (1-2). pp. 197-206.
ISSN 0378-1135 doi: 10.1016/j.vetmic.2012.05.029 Available
at <https://centaur.reading.ac.uk/28416/>

It is advisable to refer to the publisher's version if you intend to cite from the
work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1016/j.vetmic.2012.05.029>

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law,
including copyright law. Copyright and IPR is retained by the creators or other
copyright holders. Terms and conditions for use of this material are defined in
the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Accepted Manuscript

Title: *Escherichia coli* isolates from extraintestinal organs of livestock animals harbour diverse virulence genes and belong to multiple genetic lineages

Authors: Guanghai Wu, Ralf Ehricht, Muriel Mafura, Matthew Stokes, Noel Smith, Geoff C. Pritchard, Martin J. Woodward



PII: S0378-1135(12)00338-0
DOI: doi:10.1016/j.vetmic.2012.05.029
Reference: VETMIC 5780

To appear in: *VETMIC*

Received date: 13-3-2012
Revised date: 20-5-2012
Accepted date: 22-5-2012

Please cite this article as: Wu, G., Ehricht, R., Mafura, M., Stokes, M., Smith, N., Pritchard, G.C., Woodward, M.J., *Escherichia coli* isolates from extraintestinal organs of livestock animals harbour diverse virulence genes and belong to multiple genetic lineages, *Veterinary Microbiology* (2010), doi:10.1016/j.vetmic.2012.05.029

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

***Escherichia coli* isolates from extraintestinal organs of livestock animals harbour
diverse virulence genes and belong to multiple genetic lineages**

Guanghui Wu^{1*}, Ralf Ehrlich², Muriel Mafura³, Matthew, Stokes,³ Noel Smith⁴, Geoff
C. Pritchard⁵ and Martin J. Woodward^{3, 6}

¹Epidemiology, Surveillance and Risk Group, Animal Health and Veterinary
Laboratories Agency (AHVLA), New Haw, Addlestone, Surrey KT15 3NB

²Alere Technologies GmbH, Löbstedter Straße 103 -105, D-07749, Jena – Germany

³Department of Bacteriology, Animal Health and Veterinary Laboratories Agency
(AHVLA), New Haw, Addlestone, Surrey KT15 3NB

⁴Department for Bovine Tuberculosis, AHVLA, Weybridge

⁵AHVLA Bury St. Edmunds, UK

⁶ Department of Food and Nutritional Sciences, The University of Reading, PO Box
226, Whiteknights, READING, RG6 6AP

Corresponding Author:

* Dr. G. Wu: Tel: +44 1932 359475; fax: +44 1932 349 983 and email:

guanghui.wu@ahvla.gsi.gov.uk

Keywords: extraintestinal *E. coli* (ExPEC), multi-locus sequence typing (MLST),
SplitsTree, ClonalFrame, microarrays

25 Abstract

26 *Escherichia coli*, the most common cause of bacteraemia in humans in the UK, can also
 27 cause serious diseases in animals. However the population structure, virulence and
 28 antimicrobial resistance genes of those from extraintestinal organs of livestock animals are
 29 poorly characterised. The aims of this study were to investigate the diversity of these
 30 isolates from livestock animals and to understand if there was any correlation between the
 31 virulence and antimicrobial resistance genes and the genetic backbone of the bacteria and
 32 if these isolates were similar to those isolated from humans. Here 39 *E. coli* isolates from
 33 liver (n=31), spleen (n=5) and blood (n=3) of cattle (n=34), sheep (n=3), chicken (n=1)
 34 and pig (n=1) were assigned to 19 serogroups with O8 being the most common (n=7),
 35 followed by O101, O20 (both n=3) and O153 (n=2). They belong to 29 multi-locus
 36 sequence types, 20 clonal complexes with ST23 (n=7), ST10 (n=6), ST117 and ST155
 37 (both n=3) being most common and were distributed among phylogenetic group A (n=16),
 38 B1 (n=12), B2 (n=2) and D (n=9). The pattern of a subset of putative virulence genes was
 39 different in almost all isolates. No correlation between serogroups, animal hosts, MLST
 40 types, virulence and antimicrobial resistance genes was identified. The distributions of
 41 clonal complexes and virulence genes were similar to other extraintestinal or commensal
 42 *E. coli* from humans and other animals, suggesting a zoonotic potential. The diverse and
 43 various combinations of virulence genes implied that the infections were caused by
 44 different mechanisms and infection control will be challenging.

45

1. Introduction

Escherichia coli can be a commensal organism or a causative agent of diarrhoea or extra-intestinal infections-responsible for an estimated 120 million cases of community-acquired urinary tract infections (UTI) diagnosed worldwide annually. It can also cause neonatal meningitis, pneumonia and surgical site infections. The sepsis-associated mortalities due to *E. coli* are estimated at 868,000 per year globally (Russo and Johnson, 2003). In England, Wales and Northern Ireland, *E. coli* has been the most common cause of bacteraemia in humans for most years since 1990 with a year-on-year increases to 27,055 reports in 2010 (HPA). Extra-intestinal pathogenic *E. coli* (ExPEC) strains also cause mastitis, septicaemia, urogenital tract infections and sporadic abortions in cattle, pigs and sheep as well as respiratory infections and colisepticaemia in poultry that can lead to high morbidity and mortality resulting in significant economic losses (Gyles, 1994).

Many lines of evidence suggest links between human and animal extraintestinal pathogenic *E. coli* (ExPEC) (Clermont et al., 2011; Hannah et al., 2009; Johnson et al., 2001b; Johnson et al., 2007; Johnson et al., 2009; Moulin-Schouleur et al., 2007; Warren et al., 2008; Zhao et al., 2009). Although ExPEC from poultry, the avian pathogenic *E. coli* (APEC), have been studied extensively, only 13 ExPEC isolates from livestock animals (excluding poultry) were found in the *E. coli* multi-locus sequence typing (MLST) database (<http://mlst.ucc.ie/mlst/dbs/Ecoli>). Given this data gap, we aimed to characterize ExPEC strains, mostly from cattle and sheep isolated in England and Wales, in terms of their virulence, antimicrobial resistance profiles, MLST types and serogroups in order to compare with those isolates reported as causing infections in humans. A better understanding of these organisms and their distribution amongst different host species will be an important first step towards the prevention and control of infections in both humans and animals.

2. Materials and methods

2.1. Strain selection

E. coli isolates used in this study are shown in Fig. 1. All were isolated between 1999 and 2008 in England and Wales. They were a subset of those from an enhanced surveillance study. Case selection criteria, bacterial isolation, culture and strain identification methods have been described in detail previously (Hutchinson et al., 2011). Isolates associated with outbreaks where there were no clear alternative diagnoses to *E. coli* infection were included.

2.2. Multilocus sequence typing and data analysis

MLST was performed based on published method (Wirth et al., 2006). However some primers were re-designed, which gave better results for this panel of strains. Locus *adk* was amplified and sequenced with the following primer pairs: *adk_vla_F*, 5' GCAATGCGTATCATTCTGCTTGG 3' and *adk_vla_R*, 5' GGCTTGGTGCCGTCAACTTTC 3'. Locus *fumC* was amplified with primer pairs: *fumC-P1*, 5' TCACAGGTCGCCAGCGCTTC 3' and *fumC-P2*, 5' GTACGCAGCGAAAAAGATTC 3', but sequenced with *fumC_F1*, 5' TCCCGGCAGATAAGCTGTGG 3' and *fumC_R1*, 5' CAAACGGTGCACAGGTAATG 3'. Locus *gyrB* was amplified with *gyrB-P1*, 5' TCGGCGACACGGATGACGGC 3' and *gyrB-P2*, 5' ATCAGGCCTTCACGCGCATC 3', but sequenced with *gyrB_F1*, 5' ATTCCGACCGGTATTCACC 3' and *gyrB_R1*, 5' AGTACCGCCGTCACGCT 3'. Locus *icd* was amplified and sequenced with *icd-P1*, 5' ATGGAAAGTAAAGTAGTTGTTCCGGCACA 3' and *icd-P2*, 5' GGACGCAGCAGGATCTGTT 3'. Locus *mdh* were amplified and sequenced with *mdh-P1* 5' ATGAAAGTCGCAGTCCTCGGCGCTGCTGGCGG 3' and *mdh-P2*, 5'

TTAACGAACTCCTGCCCCAGAGCGATATCTTTCTT 3'. Locus *purA* was amplified with *purA_F*, 5' TCGGTAACGGTGTGTGCTG 3' and *purA_R*, 5' CATACGGTAAGCCACGCAGA 3', but sequenced with *purA_F1*, 5'GCGCTGATGAAAGAGATGAA 3'and *purA_R1*, 5' GAATTCGTTACCCTGCTTGC 3'. Locus *recA* was amplified with primers: *recA_vla-193*, 5' GGCCGTATCGTCGAAATCTA 3'and *recA-P1*, 5' CGCATTCGCTTTACCCTGACC 3' and sequenced with primers: *recA_vla_221*, 5' AATCTTCCGGTAAAACCACG 3' and *recA_val_919*, 5' CCTGACCGATCTTCTCACCT 3'. The PCR reactions were carried out for 1 cycle at 95 °C for 2 min, 30 cycles at 94 °C for 1 min, 54-64 °C for 1 min and 72 °C for 2 min, and then 1 cycle at 72 °C for 5 min. The following annealing temperatures were used: *adk*, *purA* and *recA* at 56 °C, *fumC* and *icd* at 54 °C, *gyrB* at 60 °C and *mdh* at 62-64 °C. The annealing temperatures were sometimes needed to be adjusted in order for a single band to be observed on the agarose gel before sending the products for sequencing. The sequencing data were imported to Bionumerics (5.1) and allelic numbers and MLST were assigned by submitting the results to the MLST database (<http://mlst.ucc.ie/mlst/dbs/Ecoli>).

Phylogenetic inferences about ancestral allelic profiles and strain interrelatedness were made using eBURST version 3 <http://eburst.nlst.net/> (Feil et al., 2004), SplitsTree4 <http://www.splitstree.org/> (Huson and Bryant, 2006) and ClonalFrame version 1.1 <http://www.xavierdidelot.xtreemhost.com/clonalframe.htm>. Five independent runs of Markov chain were employed and the calculated Gelman-Rubin statistics for all parameters were below 1.20, indicating satisfactory convergence between tree replicates (Didelot and Falush, 2007).

2.3. Serotyping, phylotyping, microarrays and PCR analyses of virulence and antimicrobial resistance genes

Serotyping was carried out as described earlier (Geue et al., 2010; Hutchinson et al., 2011; Wu et al., 2010b). The details of microarray analyses have been reported (Geue et al., 2010; Monecke et al., 2011) and the layout of the array can be found by following the link (http://alere-technologies.com/fileadmin/Media/Paper/Ecoli/Supplement_Geue_layout_E_coli.xlsx).

PCR virulence typing was performed according to the published method (Johnson and Stell, 2000) and following genes were detected by PCR: *bmaE*, *cvaC*, *fimH*, *focG*, *fyuA*, *gafD*, *ibeA*, *iutA*, *kpsMT* II, *kpsMT* III, *kpsMT* K1, *papC*, *papEF*, *papA*, *papG* I, *papG* II, *papG* III, *rfc*, *sfa/focD*, *traT*. The genes detected by PCR were analyzed together with microarray data with Bionumerics (5.1).

E. coli phylotyping was based on the detection of *chuA*, *yjaA* and TspE4.C2 by PCR (Clermont et al., 2000). However additional primers were used for *chuA*, 5' ATGATCATCGCGGCGTGCTG 3' and 5' AAACGCGCTCGCGCCTAAT-3'; *yjaA*, 5' TGTTCGCGATCTTGAAAGCAAACGT 3' and 5' ACCTGTGACAAACCGCCCTCA 3' and TspE4.C2 5' GCGGGTGAGACAGAAACGCG 3' and 5' TTGTCGTGAGTTGCGAACCCG 3'. PCR conditions for above primers were 1 cycle at 94°C for 4 min; 30 cycles of 94°C 30 sec, 65°C 30 sec, 72°C 30 sec; with a final extension at 72°C for 5 min.

Results and discussion

3.1 Serogrouping and phylogenetic grouping of ExPEC isolates

All isolates were from internal organs or blood of livestock animals to avoid selecting commensal *E. coli*. Only the results of isolates from different outbreaks were included here with the exception of B2710 (from liver of a calf) and B2711 (from blood of a calf) that were from the same outbreak. Isolates were from different geographic areas in England and Wales. Nineteen different serogroups were identified with O8 being the most common (n=7), followed by O101 and O20 (both n=3) and O153 (n=2), 9 were un-typable and other 15 isolates belonged to 15 different serogroups (Fig. 1).

The most common ST complexes were ST23 (n=7), ST10 (n=6), ST155 (n=3) and ST117 (n=3) (Fig. 1). Based on the information in the MLST database, clonal complexes ST10, 23, and 155 all contain multiple pathotypes of *E. coli*, including enteropathogenic *E. coli* (EPEC), verotoxigenic *E. coli* (VTEC), diffused adhesive *E. coli* (DAEC), enteroaggregative *E. coli* (EAEC), enterotoxigenic (ETEC) from humans, ExPEC strains from both animals and humans and avian pathogenic *E. coli* (APEC). Clonal complex ST155 also contains enteroinvasive (EIEC) from humans, while ST117 appears to be only associated with ExPEC strains from birds, cats, cattle and humans. Other sequence types or clonal complexes, *i.e.* ST 295, 349, 446, 469, 372 identified in this work have also been found amongst APEC and ExPEC isolates from humans and animals. In addition, the clonal complex ST69 was found only in human ExPEC and ST101 in APEC previously.

There were only 13 ExPEC from livestock animals (excluding APEC) in the database at the start of this work; strains studied here only shared the clonal complexes ST10 and 23 with them. Furthermore, a number of new types were identified in this work; therefore the clonal origins of ExPEC from livestock appeared to be very diverse.

In spite of the diversity, they were not found within the clonal complexes that are specific to diarrhoeic *E. coli*, but rather showed to be similar to ExPEC isolates from humans or animals. This result similar to the observation of the diverse clonal origins of human ExPEC isolates (Jaureguy et al., 2008), although direct comparison between these two studies cannot be made as two different MLST schemes were used.

In January 2012, there were 4245 isolates in the *E. coli* MLST database belonging to 2545 STs. The 29 STs identified in this work were compared with all identified *E. coli* MLST types using the BURST (Based upon Related Sequence Types) algorithm (Feil et al., 2004). it is clear from this analysis that these 39 ExPEC isolates were distributed widely among multiple clonal complexes. However, they were not found among ST95 complex that consists mostly of human and poultry strains nor were they in ST131 and ST73 that consists mostly of strains from humans and their pets (Fig. S1).

The results also showed that the distribution of clonal complexes was not animal specific. For example, ST88 (within the ST23 complex) contained isolates from cattle, a pig and sheep and ST783 contained isolates from both sheep and cattle. This is consistent with other results in the database (Wirth et al., 2006). Furthermore, strains of the same sequence types often belonged to different serogroups and strains of the different sequence types sometimes were of the same serogroup. Similarly, strains of the same sequence types were isolated from different internal organs and those of the different sequence types were often found in same internal organs (Fig. 1).

Phylogenetic typing showed that the isolates were from groups A (n=16), B1 (n=12), B2 (n=2) and D (n=9) (Fig. 2, 3, 4, S2 and S3). As expected ST10 complexes belonged group A; ST 101 belonged to B1; ST 106 belonged to group D, consistent with information in MLST database and published works (Okeke et al., 2010; Wirth et al.,

2006). Isolates from ST23 complexes were also assigned to group A due to the lack of amplification products despite of repeated attempts with two different primer pairs for TspE4.C2. Based on the phylogenetic groups determined by multilocus enzyme electrophoresis (MLEE), isolates of the ST23 complex may belong to either group A or group B1 (Okeke et al., 2010). It has been reported that the Clermont method (Clermont et al., 2000) used here for assigning strains to phylogenetic groups, is correct 80-85% of the time, and works best for assigning strains to groups B1 and B2 and worst with group A. This is because the assignment of group A relies on the absence of amplification products (Gordon et al., 2008).

Most studies on human ExPEC were concentrated on urinary tract infections and phylogenetic group B2 and D were predominately found (Kanamaru et al., 2006; Zhang et al., 2002). However, a more recent study on bacteraemic *E. coli* isolates from two French hospitals has shown that human ExPEC strains are highly diverse and distribute without bias into five major lineages, corresponding to the classical *E. coli* phylogroups (A+B1, B2, D and E) and group F (which contains strains previously assigned to group D) (Jauregui et al., 2008). Also, a recent study in Denmark found groups A and B1 are associated with sites of infection other than urinary tract; in that study patients with hepatobiliary septicaemia were relatively numerous (Bukh et al., 2010). As the animal isolates studied here were mainly from animal liver and most of them belonged to group A or B1, it would be interesting to determine if phylogroups A and B1 are associated with infections in the liver of animals and humans.

The MLST sequences were further analysed with SplitsTree4 to investigate the influence of recombination on the evolution of each locus. It constructs a split network connection between taxa whenever there is a phylogenetic inconsistency due to homoplasy or recombination. Recombination is generally inferred when competing

splits have equal support. Analysis of sequence data revealed extensive network structures for individual loci and the concatenated sequences (Fig. 2 and Fig. S2), implying extensive sequence exchange between lineages (recombination).

This sequence based analysis provided additional information regarding the relationship among isolates, for example, two B2 isolates 2723 ST491 and 2774 ST372 that only shared two identical loci, *gyrB* and *icd* were found to be related at the sequence levels at other 5 loci as well. The group D isolates were apparently very diverse and did not form a monophyletic group. For example, group D isolate 2721 (ST753) shared similar *fumC* sequences with ST10 isolates and *recA* was identical. In addition, group D isolates of ST117 were similar to B2 isolates of ST491 and ST372 at *adk*, *fumC*, *gyrB* and *purA* (Fig. 2 and 2S). All group D isolates of different MLST types formed distinctive long branches with strong support (bootstrap 100%, Fig. 2.). This kind of diversity was similar to group D isolates previously described for EAEC strains (Okeke et al., 2010). The two fully genome sequenced group D strains are also distinct from each other (Touchon et al., 2009). Most groups A and B1 isolates had shorter branches, suggesting that they were closely related (Fig. 2). This is consistent with previous observations that group A and B1 strains were not distinctly separated and were considered as sister groups (Jauregui et al., 2008; Lecomte et al., 1998; Okeke et al., 2010). The results suggest that the extensive recombination has obscured the phylogenetic relationships among this group of isolates.

To further explore the impact of recombination on the phylogenetic relationship of these isolates the concatenated 7 gene sequences derived from the MLST analyses were analysed with ClonalFrame, a program designed to infer the clonal relationship of bacteria and the chromosomal position of homologous recombination events that disrupt a clonal pattern of inheritance (Didelot and Falush, 2007). The results of ClonalFrame

analysis were similar to those obtained with SplitsTree (Fig. 2 & Fig. S3) and showed strong evidence of recombination. ClonalFrame is able to estimate the relative frequency of recombination compared to point mutation in genetic diversification. For our data set p/θ was estimated as 0.72 (95% CI 0.25 to 1.65). This is comparable to the ratio of 0.32-2.14 determined by Wirth ((Wirth et al., 2006) and similar to results for ST10 and triple locus variants (largely ECOR A EAEC strains) described by Okeke et al (Okeke et al., 2010). The relative impact of recombination compared to point mutation (r/m) was 1.9 (95% CI 0.83 to 3.77) which suggests that a basepair is almost twice as likely to change by recombination than by mutation. After analysing the core genes of 20 fully sequenced *E. coli*, Touchon et al (Touchon et al., 2009) also found that a gene conversion event is twice as likely as a mutation to occur at a given position. However these r/m ratios are considerably lower than previously found with EAEC strains (2.79-4.52) (Okeke et al., 2010). Touchon et al suggest that due to the short tract length of DNA (at an estimated tract length of 50 bp) involved in gene conversion, “the substantial level of gene conversion in *E. coli* did not blur the phylogenetic signal and a meaningful robust tree topology can be extracted from the sequence (Touchon et al., 2009)”. Extensive recombination among isolates of different phylogenetic groups was observed among isolates studied here. This is consistent with the higher recombination rates observed among pathogenic *E. coli* (Wirth et al., 2006).

3.2. Virulence typing results

Thirty-six different combinations of virulence gene patterns were identified among 39 isolates studied here (Fig. 3). At the 50% similarity level based on Jaccard similarity coefficient, isolates could be divided into 16 groups (with numbers per group ranging from 1-8 with a mean of 2.43, a standard deviation of 2.37, and both the mode

and median of 1), indicating remarkable diversity. All isolates harboured *fimH* (a gene for type I fimbriae). Over 60% had *traT* and *iss* (both are serum resistant genes) and *lpfA* (the gene for the long polar fimbriae). About half of isolates contained genes that are related to iron utilisation, *fyuA* (yersiniabactin) and *iutA* (aerobactin). More than 30% of isolates harboured *iroN* (enterobactin siderophore receptor protein), *mchF* and *mcmA* (both related to microcin production) and *prfB* (a fimbriae gene). More than 20% of isolates had *cvaC* (encodes Colicin V), *papC*, *bmaE* and *fli7G* (all associated with fimbriae production). The remaining genes on the arrays were found in less than 20% of isolates. *E. coli* serine protease gene *espP* normally associated with VTEC (Brunner et al., 1997) was found in 6 isolates. This gene has been found in cattle and swine F165-positive strains (Dezfulian et al., 2003). Other genes identified among this group of isolates such as *astA* (heat-stable enterotoxin), *lpfA*, *pic* (serine protease) were also found in diarrhoeic bacteria especially among EAEC as well as among UPEC strains (Abe et al., 2008). Genes such as *astA*, *iss*, *iroN*, *mchF*, *pfrB* and *cdtB* were found in *E. coli* strains from both diarrhoea and extra-intestinal infections (Wu et al., 2010b).

The numbers of virulence related genes in these isolates were from 1-18; this includes genes for microcins and iron utilization that may not be directly associated with the virulence *per se*, but confers competitive advantage for their hosts. Isolates 2774, a member of B2 group, had the highest number of putative virulence genes including those for fimbria, colicin, and iron acquisition. It was the only isolate that harboured *ibeA*, a gene also found in APEC (Germon et al., 2005). The product of *ibeA* is responsible for the invasion of brain endothelium and is associated with human neonatal meningitis (Huang et al., 2001a; Huang et al., 2001b). Isolates 2774 belonged to ST372 complex; this clonal complex has been found among APEC and ExPEC strains from humans, dogs and cats. It was isolated from the liver of a calf that died of septicaemia in

2006 together with 6 other 7 day olds calves. Database record shows that another *E. coli* isolate that was not studied here was from the brain of a calf during the same outbreak. Strain 2778 harboured 17 virulence factors and belonged to ST117. This ST is also found in APEC and human ExPEC. Three ST117 isolates (2762, 2720 and 2778) all isolated from the liver or blood of calves, all harboured *pic* and *vat*, but did not have *ctdB*, *f17AG*, *gafD*, *papG* II or *tsh*. Isolates 2762 was from a farm in Lincolnshire in 2004, 2720 was from a farm in Dorset in 2005 and 2778 was from a farm in Cumbria in 2007. The *vat* gene has been found among ST117 isolates from human patients with cirrhosis (Bert et al., 2010). Genes *vat*, *sat* and *pic* are also known to be associated with urinary tract isolates (Restieri et al., 2007).

Isolate 2750 was from the liver of a calf had *kpsMT* II genes belonged to group D and ST69 complex that is usually associated with human UTI or meningitis. It is known that ST69 UTI or bacteraemic isolates consists of 5 serogroups including O17 (Tartof et al., 2005), the serogroup that 2750 belonged to. Two other isolates, 2710 and 2711, were from the same farm and both belonged to ST10 and serogroup O101. Isolate 2710 was from the liver of a calf harboured more virulence and antimicrobial resistance genes than isolate 2711 that was isolated from blood of a calf, suggesting the acquisition or loss of virulence factors either during the infection or after the isolation. Isolate 2839 that was from the liver of a calf contained 15 virulence and 8 antimicrobial resistance genes including a gene encoding extended spectrum β -lactamase (*bla*_{CTX-M-group-9}) (See later section). No virulence factors (apart from *fimH* that can be found in both commensals as well as pathogens) were detected in some liver isolates, which suggests the presence of unknown or untested virulence factors. This work is consistent with early findings that described the overlapping virulence genes from *E. coli* isolated from diseased calves and pigs and humans (Dezfulian et al., 2003; Girardeau et al., 2003).

Different virulence genes were identified among these ExPEC isolates, which indicated the presence of possible subtypes of ExPEC strains. For *E. coli* causing enteric/diarrhoeal diseases, at least six pathotypes with specific virulence genes have been described (Nataro and Kaper, 1998). The research on the pathogenic mechanisms of ExPEC strains is lagging behind, but nevertheless it has been recognised that different mechanisms must exist among ExPEC to cause diseases (Brzuszkiewicz et al., 2006; Johnson et al., 2001a). Marrs *et al.* have started to define these different ‘subtypes’ but these authors have pointed out the need for more detailed studies to define specific virulence genes in these organisms (Marrs et al., 2005).

3.3. Antimicrobial resistance gene carriage

The isolates used in this work were selected on the basis of either association with or, probable cause of, diseases and not for antimicrobial resistance. Twenty-one (53%) isolates harboured at least one antimicrobial resistance genes and some have up to 9 antimicrobial resistance genes (Fig. 4). The most common antimicrobial resistance genes were *strB*, *bla*_{TEM} and *sul2* that were found in more than 35% of the isolates. The *aadA1*, *sul1* and *aphA* gene were in more than 10% of isolates. A *bla*_{CTX-M-group-9} gene was found in isolate 2839 from the liver of cattle. The *intI* 1 gene (for type I integron) was found to be closely associated with *sul1* and *aadA1*, so was *aphA* with *sul2* and *bla*_{TEM-1} with *strB*.

It has been suggested that there may be a fitness ‘trade-off’ for organisms between virulence and resistance. For example, B2 strains are considered to be more virulent (contain P fimbriae and are α , β -haemolytic) but less resistant to antimicrobials than B1 (Bukh et al., 2010; Jaureguiy et al., 2007; Johnson et al., 1991; Johnson et al., 1994). In Denmark, the prevalence of antibiotic resistance harboured by phylogroup

decreases in the order of group D>A> B1> B2 (Bukh et al., 2010). In this work, the P-fimbriae positive strains harboured anything from 0 to 9 antimicrobial resistance genes. Resistance genes were found in 6/9 (67%) of group D strains, 12/16 (75%) group A strains, 4/12 (33%) group B1 strains. No resistance genes were identified in two group B2 strains. The average number of antimicrobial resistance genes in group D strains was 7.8; group A was 6.8 and group B1 was 3.8.

3.4. Phylogenetic distribution of virulence genes and antimicrobial resistance genes

There was no correlation between the carriage of virulence and antimicrobial resistance genes and the genetic backbone of strains (Fig. 3 and 4). For example, a variety of virulence genes were found in ST10, ST117 and ST23 strains. Occasionally isolates of the same ST from different sources shared similar virulence genes. Furthermore, isolate 2770 from the liver of a sheep and 2839 from the liver of a calf both belonged to ST783 and shared 13 virulence genes; but the sheep isolate harboured *iha* and *bmaE*, while cattle isolate had *ireA* and *papA*. Similarly, isolates that were of the same sequence type contained different antimicrobial resistance genes.

Many virulence and antimicrobial resistance genes identified in this set of isolates have disseminated into multiple genetic backgrounds. Previous studies have revealed the parallel evolution of EPEC with multiple acquisitions of virulence genes in different background (Lacher et al., 2007; Reid et al., 2000; Wirth et al., 2006). After studying multiple pathotypes of *E. coli*, Escobar-Paramo et al concluded that a specific genetic background is required for acquisition and expression of virulence factors in *E. coli* and the ExPEC associated virulence genes were linked to the phylogenetic group B2 in that study (Escobar-Paramo et al., 2004).

Although some sequence types of *E. coli* such as ST10 and ST155 contain multiple pathotypes, others appear to be pathotype specific. For example ST69 is predominately associated with UPEC isolates from human and ST11 contained only EHEC O157:H7 or EPEC O55:H7 strains. Therefore, the interplay between horizontally transferred genetic elements and genetic backbone of bacteria is rather complicated

Two major clonal complexes found among these isolates are ST10 and ST23 and both harboured various combinations of virulence genes. Generally speaking, more virulence genes were associated with ST23 complex than with ST10. As all isolates were from diseased animals, this discrepancy in the number virulence genes was probably due to our limited understanding of those organisms. Many more virulence factors are yet to be discovered; only then we will be able to understand better the association between virulence genes and genetic backbones of *E. coli*. Considerable diversity was observed among strains causing similar infections in animals. Perhaps this is not surprising as even among the isolates of the same serotypes, differences can be substantial due to the constant acquisition and loss of genes (Wu et al., 2010a; Wu et al., 2008). Further work is needed to understand the mechanisms of pathogenicity and their potentials in causing diseases in different hosts.

4. Conclusion

E. coli isolates from extra-intestinal organs of livestock animals belonged to multiple serogroups and phylogenetic groups. The clonal origins and the virulence genes harboured by these strains were similar to ExPEC from humans and other animals, suggesting the zoonotic potential. The many different combinations of virulence factors indicated multiple pathogenic mechanisms. No correlation was found between the genetic backbone, and the virulence and antimicrobial resistance gene content. Because

of this remarkable genetic diversity, it will be challenging to control the infections caused by ExPEC.

Acknowledgements

This work was supported by Defra through the VLA Seedcorn funding SC218, the Endemic Diseases and Welfare programme, and the non-statutory zoonoses project (FZ2100). We thank Dr. Xavier Didelot for helping with the analysis of MLST data by ClonalFrame, Michaela Williams from the Health Protection Agency for designing the extra phylotyping primers and Katharine Lynch for helping with serotyping. We thank Sarah Brown for helping with preparing the figures.

Conflict of interest

There is no conflict of interest.

References:

- Abe, C.M., Salvador, F.A., Falsetti, I.N., Vieira, M.A., Blanco, J., Blanco, J.E., Blanco, M., Machado, A.M., Elias, W.P., Hernandes, R.T., Gomes, T.A., 2008. Uropathogenic *Escherichia coli* (UPEC) strains may carry virulence properties of diarrhoeagenic *E. coli*. FEMS Immunol Med Microbiol 52, 397-406.
- Bert, F., Johnson, J.R., Ouattara, B., Leflon-Guibout, V., Johnston, B., Marcon, E., Valla, D., Moreau, R., Nicolas-Chanoine, M.H., 2010. Genetic diversity and virulence profiles of *Escherichia coli* isolates causing spontaneous bacterial peritonitis and bacteremia in patients with cirrhosis. J Clin Microbiol 48, 2709-2714.

- 414 Brunder, W., Schmidt, H., Karch, H., 1997. EspP, a novel extracellular serine protease
415 of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation
416 factor V. Mol Microbiol 24, 767-778.
- 417 Brzuszkiewicz, E., Bruggemann, H., Liesegang, H., Emmerth, M., Olschlager, T., Nagy,
418 G., Albermann, K., Wagner, C., Buchrieser, C., Emody, L., Gottschalk, G.,
419 Hacker, J., Dobrindt, U., 2006. How to become a uropathogen: comparative
420 genomic analysis of extraintestinal pathogenic *Escherichia coli* strains. Proc Natl
421 Acad Sci U S A 103, 12879-12884.
- 422 Bukh, A.S., Schonheyder, H.C., Emmersen, J.M.C., Sogaard, M., Bastholm, S., Rosiev,
423 P., 2010. *Escherichia coli* phylogenetic groups are associated with site of
424 infection and level of antibiotic resistance in community-acquired bacteraemia: a
425 10 year population-based study in Denmark Journal of Antimicrobial
426 Chemotherapy 64, 163-168.
- 427 Clermont, O., Bonacorsi, S., Bingen, E., 2000. Rapid and simple determination of the
428 *Escherichia coli* phylogenetic group. Appl Environ Microbiol 66, 4555-4558.
- 429 Clermont, O., Olier, M., Hoede, C., Diancourt, L., Brisse, S., Keroudean, M., Glodt, J.,
430 Picard, B., Oswald, E., Denamur, E., 2011. Animal and human pathogenic
431 *Escherichia coli* strains share common genetic backgrounds. Infect Genet Evol
432 11, 654-662.
- 433 Dezfulian, H., Batisson, I., Fairbrother, J.M., Lau, P.C., Nassar, A., Szatmari, G., Harel,
434 J., 2003. Presence and characterization of extraintestinal pathogenic *Escherichia*
435 *coli* virulence genes in F165-positive *E. coli* strains isolated from diseased calves
436 and pigs. J Clin Microbiol 41, 1375-1385.
- 437 Didelot, X., Falush, D., 2007. Inference of bacterial microevolution using multilocus
438 sequence data. Genetics 175, 1251-1266.

- 439 Escobar-Paramo, P., Clermont, O., Blanc-Potard, A.B., Bui, H., Le Bouguenec, C.,
 440 Denamur, E., 2004. A specific genetic background is required for acquisition and
 441 expression of virulence factors in *Escherichia coli*. *Mol Biol Evol* 21, 1085-
 442 1094.
- 443 Feil, E.J., Li, B.C., Aanensen, D.M., Hanage, W.P., Spratt, B.G., 2004. eBURST:
 444 inferring patterns of evolutionary descent among clusters of related bacterial
 445 genotypes from multilocus sequence typing data. *J Bacteriol* 186, 1518-1530.
- 446 Germon, P., Chen, Y.H., He, L., Blanco, J.E., Bree, A., Schouler, C., Huang, S.H.,
 447 Moulin-Schouleur, M., 2005. *ibeA*, a virulence factor of avian pathogenic
 448 *Escherichia coli*. *Microbiology* 151, 1179-1186.
- 449 Geue, L., Schares, S., Mintel, B., Conraths, F.J., Muller, E., Ehricht, R., 2010. Rapid
 450 Microarray-Based Genotyping of Enterohemorrhagic *Escherichia coli* Serotype
 451 O156:H25/H-/Hnt Isolates from Cattle and Clonal Relationship Analysis. *Appl*
 452 *Environ Microbiol* 76, 5510-5519.
- 453 Girardeau, J.P., Lalioui, L., Said, A.M., De Champs, C., Le Bouguenec, C., 2003.
 454 Extended virulence genotype of pathogenic *Escherichia coli* isolates carrying the
 455 *afa-8* operon: evidence of similarities between isolates from humans and animals
 456 with extraintestinal infections. *J Clin Microbiol* 41, 218-226.
- 457 Gordon, D.M., Clermont, O., Tolley, H., Denamur, E., 2008. Assigning *Escherichia coli*
 458 strains to phylogenetic groups: multi-locus sequence typing versus the PCR
 459 triplex method. *Environ Microbiol* 10, 2484-2496.
- 460 Gyles, C.L. 1994. *Escherichia coli* in domestic animals and humans.
- 461 Hannah, E.L., Johnson, J.R., Angulo, F., Haddadin, B., Williamson, J., Samore, M.H.,
 462 2009. Molecular analysis of antimicrobial-susceptible and -resistant *Escherichia*

- 463 *coli* from retail meats and human stool and clinical specimens in a rural
 464 community setting. Foodborne Pathog Dis 6, 285-295.
- 465 HPA, http://www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1296686942137 accessed
 466 on 23rd sept. 2011.
- 467 Huang, S.H., Chen, Y.H., Kong, G., Chen, S.H., Besemer, J., Borodovsky, M., Jong, A.,
 468 2001a. A novel genetic island of meningitic *Escherichia coli* K1 containing the
 469 *ibeA* invasion gene (GimA): functional annotation and carbon-source-regulated
 470 invasion of human brain microvascular endothelial cells. Funct Integr Genomics
 471 1, 312-322.
- 472 Huang, S.H., Wan, Z.S., Chen, Y.H., Jong, A.Y., Kim, K.S., 2001b. Further
 473 characterization of *Escherichia coli* brain microvascular endothelial cell invasion
 474 gene *ibeA* by deletion, complementation, and protein expression. J Infect Dis
 475 183, 1071-1078.
- 476 Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary
 477 studies. Mol Biol Evol 23, 254-267.
- 478 Hutchinson, J.P., Cheney, T.E., Smith, R.P., Lynch, K., Pritchard, G.C., 2011.
 479 Verocytotoxin-producing and attaching and effacing activity of *Escherichia coli*
 480 isolated from diseased farm livestock. Vet Rec 168, 536.
- 481 Jaureguy, F., Carbonnelle, E., Bonacorsi, S., Clec'h, C., Casassus, P., Bingen, E., Picard,
 482 B., Nassif, X., Lortholary, O., 2007. Host and bacterial determinants of initial
 483 severity and outcome of *Escherichia coli* sepsis. Clin Microbiol Infect 13, 854-
 484 862.
- 485 Jaureguy, F., Landraud, L., Passet, V., Diancourt, L., Frapy, E., Guigon, G.,
 486 Carbonnelle, E., Lortholary, O., Clermont, O., Denamur, E., Picard, B., Nassif,

- 487 X., Brisse, S., 2008. Phylogenetic and genomic diversity of human bacteremic
488 *Escherichia coli* strains. BMC Genomics 9, 560.
- 489 Johnson, J.R., Goullet, P., Picard, B., Moseley, S.L., Roberts, P.L., Stamm, W.E., 1991.
490 Association of carboxylesterase B electrophoretic pattern with presence and
491 expression of urovirulence factor determinants and antimicrobial resistance
492 among strains of *Escherichia coli* that cause urosepsis. Infect Immun 59, 2311-
493 2315.
- 494 Johnson, J.R., O'Bryan, T.T., Kuskowski, M., Maslow, J.N., 2001a. Ongoing horizontal
495 and vertical transmission of virulence genes and papA alleles among *Escherichia*
496 *coli* blood isolates from patients with diverse-source bacteremia. Infect Immun
497 69, 5363-5374.
- 498 Johnson, J.R., Orskov, I., Orskov, F., Goullet, P., Picard, B., Moseley, S.L., Roberts,
499 P.L., Stamm, W.E., 1994. O, K, and H antigens predict virulence factors,
500 carboxylesterase B pattern, antimicrobial resistance, and host compromise
501 among *Escherichia coli* strains causing urosepsis. J Infect Dis 169, 119-126.
- 502 Johnson, J.R., Stell, A.L., 2000. Extended virulence genotypes of *Escherichia coli*
503 strains from patients with urosepsis in relation to phylogeny and host
504 compromise. J Infect Dis 181, 261-272.
- 505 Johnson, J.R., Stell, A.L., Delavari, P., Murray, A.C., Kuskowski, M., Gastra, W.,
506 2001b. Phylogenetic and pathotypic similarities between *Escherichia coli*
507 isolates from urinary tract infections in dogs and extraintestinal infections in
508 humans. J Infect Dis 183, 897-906.
- 509 Johnson, T.J., Kariyawasam, S., Wannemuehler, Y., Mangiamale, P., Johnson, S.J.,
510 Doetkott, C., Skyberg, J.A., Lynne, A.M., Johnson, J.R., Nolan, L.K., 2007. The
511 genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares

- 512 strong similarities with human extraintestinal pathogenic *E. coli* genomes. J
 513 Bacteriol 189, 3228-3236.
- 514 Johnson, T.J., Logue, C.M., Wannemuehler, Y., Kariyawasam, S., Doetkott, C.,
 515 DebRoy, C., White, D.G., Nolan, L.K., 2009. Examination of the source and
 516 extended virulence genotypes of *Escherichia coli* contaminating retail poultry
 517 meat. Foodborne Pathog Dis 6, 657-667.
- 518 Kanamaru, S., Kurazono, H., Nakano, M., Terai, A., Ogawa, O., Yamamoto, S., 2006.
 519 Subtyping of uropathogenic *Escherichia coli* according to the pathogenicity
 520 island encoding uropathogenic-specific protein: comparison with phylogenetic
 521 groups. Int J Urol 13, 754-760.
- 522 Lacher, D.W., Steinsland, H., Blank, T.E., Donnenberg, M.S., Whittam, T.S., 2007.
 523 Molecular evolution of typical enteropathogenic *Escherichia coli*: clonal analysis
 524 by multilocus sequence typing and virulence gene allelic profiling. J Bacteriol
 525 189, 342-350.
- 526 Lecointre, G., Rachdi, L., Darlu, P., Denamur, E., 1998. *Escherichia coli* molecular
 527 phylogeny using the incongruence length difference test. Mol Biol Evol 15,
 528 1685-1695.
- 529 Marrs, C.F., Zhang, L., Foxman, B., 2005. *Escherichia coli* mediated urinary tract
 530 infections: are there distinct uropathogenic *E. coli* (UPEC) pathotypes? FEMS
 531 Microbiol Lett 252, 183-190.
- 532 Monecke, S., Mariani-Kurkdjian, P., Bingen, E., Weill, F.X., Balière, C., Slickers, P.,
 533 Ehricht, R., 2011. Presence of Enterohemorrhagic *Escherichia coli*
 534 ST678/O104:H4 in France prior to 2011. Appl. Environ. Microbiol. Accepted.
- 535 Moulin-Schouleur, M., Reperant, M., Laurent, S., Bree, A., Mignon-Grasteau, S.,
 536 Germon, P., Rasschaert, D., Schouler, C., 2007. Extraintestinal pathogenic

- 537 *Escherichia coli* strains of avian and human origin: link between phylogenetic
538 relationships and common virulence patterns. J Clin Microbiol 45, 3366-3376.
- 539 Nataro, J.P., Kaper, J.B., 1998. Diarrheagenic *Escherichia coli*. Clin Microbiol Rev 11,
540 142-201.
- 541 Okeke, I.N., Wallace-Gadsden, F., Simons, H.R., Matthews, N., Labar, A.S., Hwang, J.,
542 Wain, J., 2010. Multi-locus sequence typing of enteroaggregative *Escherichia*
543 *coli* isolates from Nigerian children uncovers multiple lineages. PLoS One 5,
544 e14093.
- 545 Reid, S.D., Herbelin, C.J., Bumbaugh, A.C., Selander, R.K., Whittam, T.S., 2000.
546 Parallel evolution of virulence in pathogenic *Escherichia coli*. Nature 406, 64-
547 67.
- 548 Restieri, C., Garriss, G., Locas, M.C., Dozois, C.M., 2007. Autotransporter-encoding
549 sequences are phylogenetically distributed among *Escherichia coli* clinical
550 isolates and reference strains. Appl Environ Microbiol 73, 1553-1562.
- 551 Russo, T.A., Johnson, J.R., 2003. Medical and economic impact of extraintestinal
552 infections due to *Escherichia coli*: focus on an increasingly important endemic
553 problem. Microbes Infect 5, 449-456.
- 554 Tartof, S.Y., Solberg, O.D., Manges, A.R., Riley, L.W., 2005. Analysis of a
555 uropathogenic *Escherichia coli* clonal group by multilocus sequence typing. J
556 Clin Microbiol 43, 5860-5864.
- 557 Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E.,
558 Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiapello, H., Clermont,
559 O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., Karoui, M.E., Frapy, E.,
560 Garry, L., Ghigo, J.M., Gilles, A.M., Johnson, J., Le Bouguenec, C., Lescat, M.,
561 Mangenot, S., Martinez-Jehanne, V., Matic, I., Nassif, X., Oztas, S., Petit, M.A.,

- 562 Pichon, C., Rouy, Z., Ruf, C.S., Schneider, D., Tournet, J., Vacherie, B.,
 563 Vallenet, D., Medigue, C., Rocha, E.P., Denamur, E., 2009. Organised genome
 564 dynamics in the *Escherichia coli* species results in highly diverse adaptive paths.
 565 PLoS Genet 5, e1000344.
- 566 Warren, R.E., Ensor, V.M., O'Neill, P., Butler, V., Taylor, J., Nye, K., Harvey, M.,
 567 Livermore, D.M., Woodford, N., Hawkey, P.M., 2008. Imported chicken meat as
 568 a potential source of quinolone-resistant *Escherichia coli* producing extended-
 569 spectrum beta-lactamases in the UK. J Antimicrob Chemother 61, 504-508.
- 570 Wirth, T., Falush, D., Lan, R., Colles, F., Mensa, P., Wieler, L.H., Karch, H., Reeves,
 571 P.R., Maiden, M.C., Ochman, H., Achtman, M., 2006. Sex and virulence in
 572 *Escherichia coli*: an evolutionary perspective. Mol Microbiol 60, 1136-1151.
- 573 Wu, G., AbuOun, M., Hackl, E., La Ragione, R.M., Fookes, M., Fenner, J., Pan, Z.,
 574 Wenzl, P., Anjum, M.F., Woodward, M.J., 2010a. Epidemic multidrug-resistant
 575 (MDR-AmpC) *Salmonella enterica* serovar Newport strains contain three phage
 576 regions and a MDR resistance plasmid. Environmental Microbiology Reports 2,
 577 228-235.
- 578 Wu, G., Carter, B., Mafura, M., Liebana, E., Woodward, M.J., Anjum, M.F., 2008.
 579 Genetic diversity among *Escherichia coli* O157:H7 isolates and identification of
 580 genes linked to human infections. Infect Immun 76, 845-856.
- 581 Wu, G., Mafura, M., Carter, B., Lynch, K., Anjum, M.F., Woodward, M.J., Pritchard,
 582 G.C., 2010b. Genes associated with *Escherichia coli* isolates from calves with
 583 diarrhoea and/or septicaemia. Vet Rec 166, 691-692.
- 584 Zhang, L., Foxman, B., Marrs, C., 2002. Both urinary and rectal *Escherichia coli*
 585 isolates are dominated by strains of phylogenetic group B2. J Clin Microbiol 40,
 586 3951-3955.

587 Zhao, L., Gao, S., Huan, H., Xu, X., Zhu, X., Yang, W., Gao, Q., Liu, X., 2009.
588 Comparison of virulence factors and expression of specific genes between
589 uropathogenic *Escherichia coli* and avian pathogenic *E. coli* in a murine urinary
590 tract infection model and a chicken challenge model. Microbiology 155, 1634-
591 1644.
592
593
594

Figure legends:

Fig. 1. The isolates were clustered based on the allelic numbers of 7 house-keeping genes using categorical coefficient and unweighted pair group method with arithmetic averages (UPGMA, Bionumerics 5.1). Sources and serogroups of the isolates are shown.

Fig. 2. The concatenated 7 house-keeping gene sequences were analysed with SplitsTree4. Bootstrap scores greater than 90 are given at each node. The scale of the network, MLST types and phylogenetic groups of isolates were indicated.

Fig. 3. Isolates were clustered based on their virulence gene content using Jaccard coefficient and UPGMA (Bionumerics 5.1). MLST types and phylogenetic groups of isolates are shown.

Fig. 4. Isolates were clustered based on their antimicrobial resistance genes using Jaccard coefficient and UPGMA (Bionumerics 5.1). MLST types and phylogenetic groups of isolates are shown.

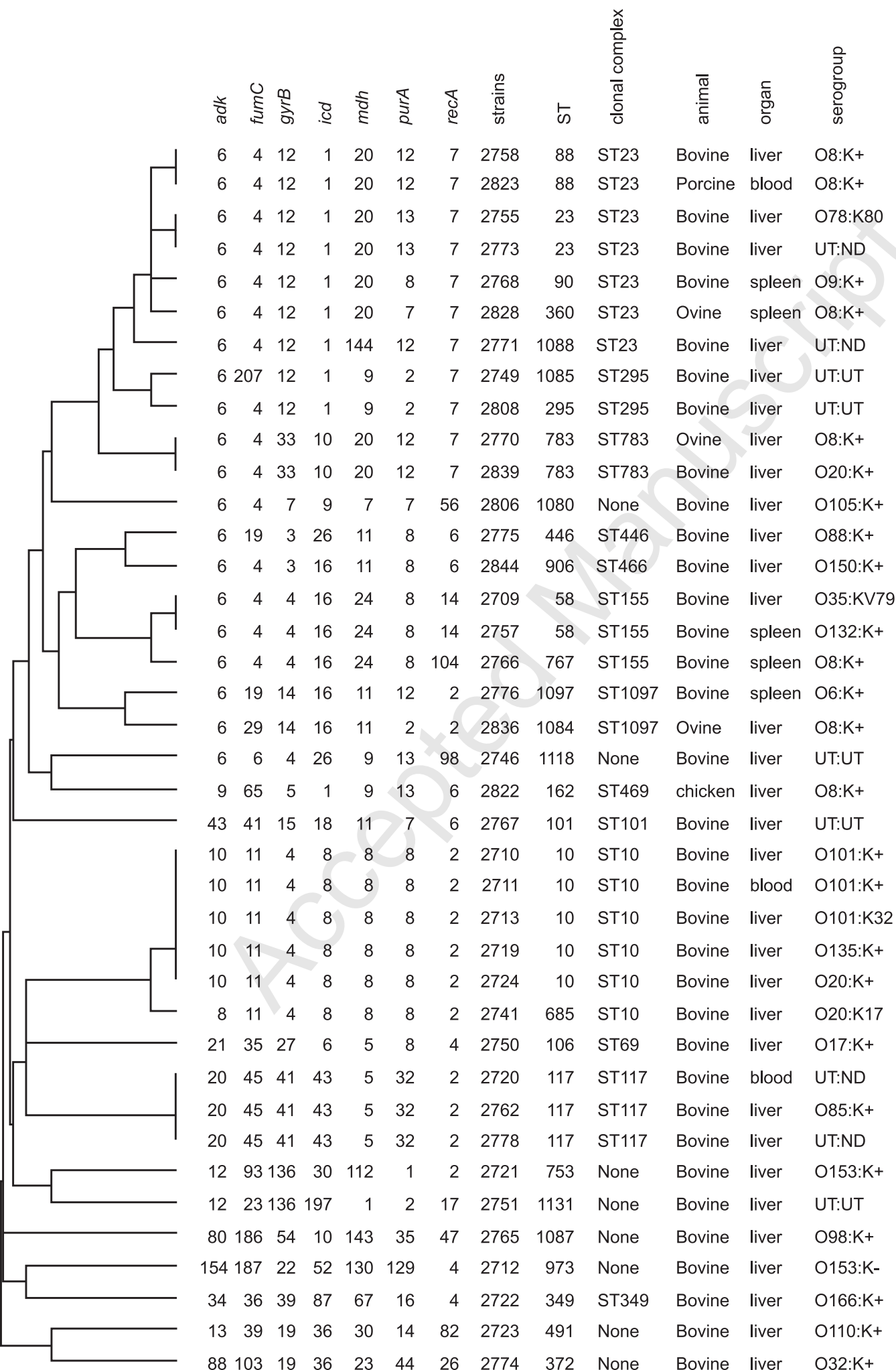


Figure 4

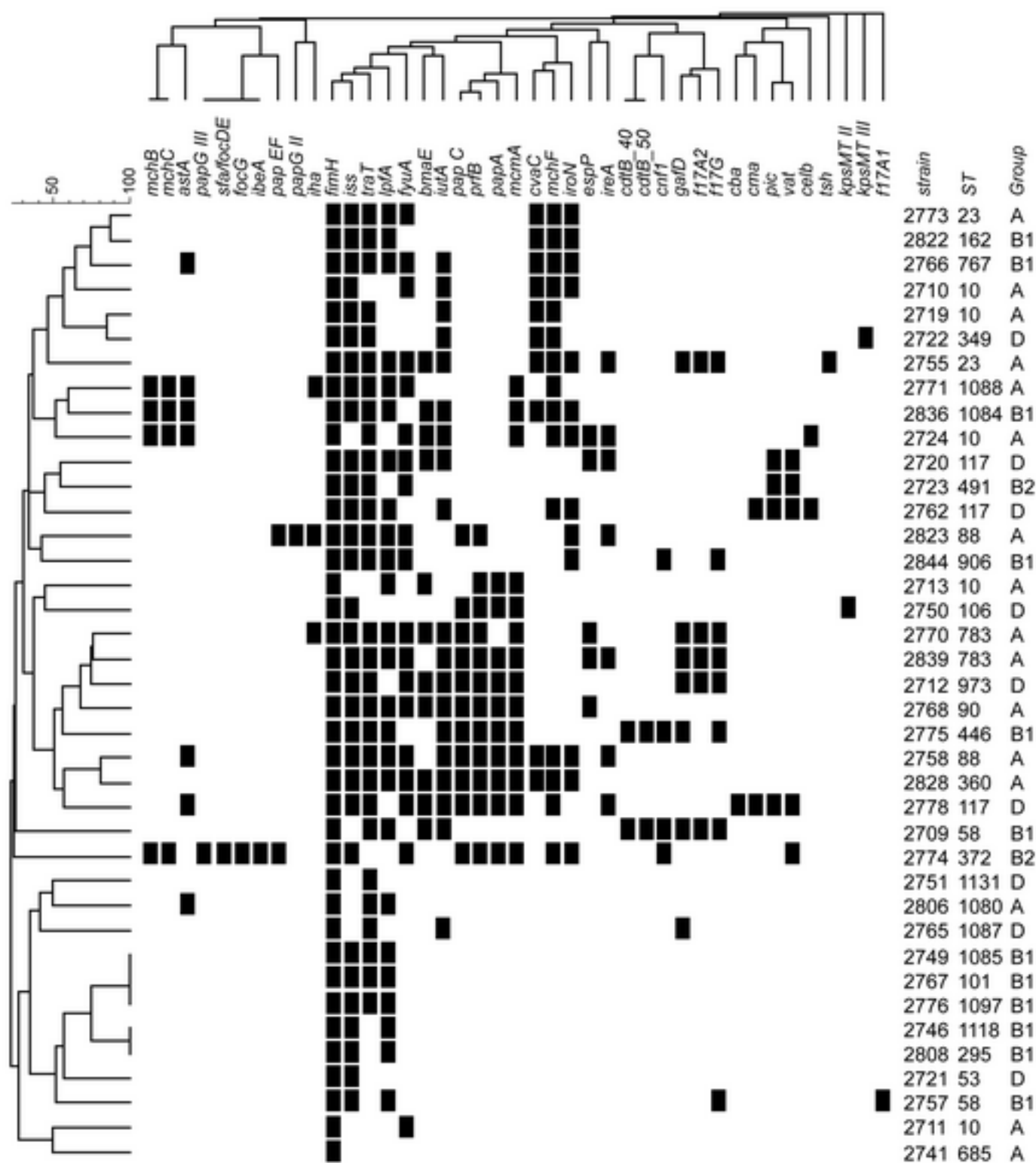


Figure 5

