The genomic epidemiology of Escherichia albertii

Rebecca Bengtsson
University of Liverpool
Kate Baker (kbaker@liverpool.ac.uk)
University of Liverpool https://orcid.org/0000-0001-5850-1949
Andrew Cunningham
Zoological Society of London https://orcid.org/0000-0002-3543-6504
David Greig
Gastrointestinal and Food Safety (One Health) Division
Shinto John
Institute of Zoology
Shaheed Macgregor
Wildlife Health Services
Katharina Seilern-Moy
Wildlife Health Services
Simon Spiro
Wildlife Health Services
Charlotte Chong
Clinical Infection, Microbiology, and Immunology
Malaka De Silva
Clinical Infection, Microbiology, and Immunology
Claire Jenkins
Gastrointestinal and Food Safety (One Health) Division
Becki Lawson
Institute of Zoology

Article

Keywords:

Posted Date: September 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2035146/v1

Version of Record: A version of this preprint was published at Nature Communications on March 27th, 2023. See the published version at https://doi.org/10.1038/s41467-023-37312-3.

Abstract

Escherichia albertii is a recently identified gastrointestinal bacterial pathogen of humans and animals which is typically misidentified and generally only detected during genomic surveillance of other Enterobacteriaceae. The incidence of *E. albertii* is likely underestimated and its epidemiology and clinical relevance are poorly characterised. Here, we whole genome sequenced *E. albertii* isolates from humans (n = 83) and birds (n = 79) in Great Britain and analysed a broader public dataset (n = 475) to address these gaps. We found human and avian isolates typically (90%; 148/164) belonged to host-associated monophyletic groups with distinct virulence and antimicrobial resistance profiles. Overlaid patient epidemiological data suggested that human infection was likely related to travel and possibly foodborne transmission. The Shiga toxin encoding *stx2f* gene was associated with clinical disease (OR = 10.27, 95% CI = 2.98–35.45 p = 0.0002) in finches. Our results suggest that improved future surveillance will further elucidate disease ecology and public and animal health risks associated with *E. albertii*.

Introduction

Escherichia albertii, a Gram-negative gastrointestinal pathogen of humans and animals, was first confirmed as a novel bacterium in 2003 (1–3). This bacteria is often mis-identified because it is difficult to differentiate from *Shigella* species and pathotypes of diarrhoeagenic *E. coli* (DEC) using phenotypic tests (4, 5). Implementation of PCR for the detection of a wide range of gastrointestinal (GI) pathogens including DEC, and the use of whole genome sequencing (WGS) for identification and typing, has provided a more robust and reliable approach for the identification and characterisation of this pathogen (6–8). As a result of the implementation of routine WGS at the United Kingdom Health Security Agency (UKHSA, formerly Public Health England), there has been an increase in the number of detections of *E. albertii* in individuals captured by microbiological surveillance (UKHSA, unpublished data, reported herein).

Although detection and speciation prior to the genomic era was challenging, the pathogenic traits of *E. albertii* are well described (4). Like certain DEC pathotypes, specifically the enteropathogenic *E. coli* (EPEC) and a subset of Shiga toxin-producing *E. coli* (STEC), the genome of *E. albertii* contains the locus of enterocyte effacement (LEE) pathogenicity island encoding a type III secretion system involved in attachment of the pathogen to the gut mucosa (9, 10). Colonisation of EPEC and *eae* gene positive (a marker of LEE) STEC in both humans and animals can lead to the formation of attaching and effacing (A/E) lesions on the intestinal epithelial cells (11). Cytolethal distending toxin (CDT) is encoded by the *cdtABC* operon and is classified into five subtypes based on sequence variation of the *cdtB* gene (*cdtB*-I to *cdtB*-V). Of these, *cdtB* subtypes I/II/III/V have been identified in *E. albertii* (10, 12). The *stx* gene encoding for Shiga toxins, predominantly the *stx2f* subtype, has been found in certain strains of *E. albertii* (13). Although these virulence determinants are well described in *E. albertii*, their distribution and clinical relevance across species requires further elucidation.

Clinical symptoms in human patients caused by *E. albertii* infection are similar to those caused by EPEC and typically include watery diarrhoea, dehydration, abdominal pain, vomiting and fever (13, 14). Over the last decade, outbreaks of GI disease in people in Japan have been attributed to *E. albertii* following re-examination of the original microbiological findings using genomic typing methods, such as multilocus sequence typing (MLST) (15, 16). However, because of the challenges around detection and identification which have hampered systematic surveillance, data on the epidemiology, source and transmission routes of *E. albertii* infections are sparse. In England, to date commercial GI PCR panels have been adopted by approximately 25% of diagnostic microbiology laboratories in the National Health Service network (17). Furthermore, not all the commercial GI PCR panels target *eae*, and not all diagnostic laboratories refer samples to the Gastrointestinal Bacteria Reference Unit (GBRU) at UKHSA for further identification. These limitations of the current surveillance mechanisms likely result in a considerable under ascertainment of cases, and the true burden of human infection caused by *E. albertii* remains unknown.

In addition to infecting people, *E. albertii* can infect birds and other animals, in which the prevalence and pathogenicity is also unclear. In the mid-1990s, multiple mortality incidents of Fringillidae (finch) species were observed in Scotland with a bacterium, later identified as *E. albertii*, hypothesised to be the cause of death (2, 18, 19). Similarly, in 2004, large-scale mortality of a finch species (*Carduelis flammea*) occurred in Alaska, United States of America (USA), with *E. albertii* as the probable aetiology (2). Active molecular surveillance studies for *E. albertii* have since detected the bacterium in dead and apparently healthy birds of multiple orders and species from Australia, Asia and North America (2, 20–22). *Escherichia albertii* also has been detected in poultry faeces/GI tract contents and meat (e.g. (8, 23–25) and in domestic mammals (e.g. pig, cat) and both terrestrial and marine wild mammal species (e.g. raccoon, seal, bat) (8, 26, 27). Although the occurrence and significance to mammal host health remains uncertain, there is a growing body of evidence that avian hosts may act as a reservoir of infection (21, 25). Thus, the extent of associated disease in birds and the relationship of bird and human infections requires further investigation.

Here, we performed WGS analysis on *E. albertii* isolates from humans and birds in Great Britain (GB) from archives held at the UKHSA and the Zoological Society of London (ZSL), respectively, to investigate the epidemiology of this recently identified pathogen. The aims of the study were to integrate the phylogenetic and epidemiological data in order to gain insights into the ecology of *E. albertii* among people and birds, to better understand the risk factors (e.g. recent international travel) associated with human infection, and to infer the likely significance of *E. albertii* infection to avian host health. Owing to the relative importance of *Enterobacteriaceae* as a reservoir for antimicrobial resistance (AMR) genes, we also describe and compare AMR profiles recovered from the two host groups.

Materials And Methods

Human isolates and epidemiological data collection

Diagnostic algorithms for the detection of *E. albertii* are not included in the UK Standard Microbiology Investigation of Gastroenteritis protocols used by local hospital diagnostic laboratories (https://www.gov.uk/government/publications/smi-s-7-gastroenteritis-and-diarrhoea). Between 2014 and 2021, isolates from faecal specimens from hospitalised cases or cases in the community, were either submitted to the GBRU at UKHSA from local hospital diagnostic laboratories

in England having been mis-identified as *Shigella* species or DEC, or were cultured from faecal specimens sent to GBRU for molecular testing. At GBRU, bacteria cultured from faecal specimens on MacConkey agar following aerobic incubation overnight, were tested for virulence genes that define the different pathotypes of DEC using PCR, including *eae* which is a characteristic of EPEC, STEC and *E. albertii* (7).

All *eae*-positive isolates were genome sequenced and bacterial identification was confirmed from the genome using a kmer-based approach, as described previously (28). In total, all 83 isolates identified as *E. albertii* using this approach were included in this study (Supplementary Table 1). Where available, human isolates were linked to demographic data, including age category, gender, and travel history (Supplementary table 1).

Bird isolates and epidemiological data collection

Wild bird derived *E. albertii* isolates (n = 74) were obtained through scanning surveillance of dead wild birds conducted by ZSL over the period 2000–2019 inclusive (Supplementary table 2). Members of the public reported observations of wild bird mortality, typically in the vicinity of garden bird feeding stations. Carcasses were submitted from a subset of mortality incidents for post-mortem examination. Coverage was across Great Britain, although the majority of wild bird submissions and those from which *E. albertii* was isolated were from England (England n = 63 isolates, Wales n = 6, Scotland n = 5). Post-mortem investigations were conducted following a standardised protocol, supported by parasitological and microbiological examination as routine, combined with histological examination and other ancillary diagnostic testing as indicated based on macroscopic abnormalities. Liver and small intestinal tract contents were routinely sampled for microbiological examination using a standardised protocol (29). Semitranslucent, butyrous, non-lactose fermenting and oxidase negative colonies of Gram-negative rods to coccobacilli were subjected to an Analytical Profile Index 20E biochemical test (bioMerieux): isolates tentatively identified as *E. albertii* were cryo-archived at -80 degrees C. Where *E. albertii* was isolated from multiple wild birds examined from the same mortality incident, a single isolate was submitted to GBRU with two exceptions where two isolates were typed. An available archive of similarly identified *E. albertii* isolates from clinical examinations (n = 2) and post-mortem examinations (n = 3) of captive birds in the zoological collection at ZSL was also included (Supplementary Table 2). Additionally, a single *E. albertii* isolate was identified from a sample of small intestinal tract contents collected from a dead wild bird examined post-mortem using the UKHSA diagnostic algorithm for human faecal samples.

The inferred significance of *E. albertii* infection to wild and captive zoo bird health (i.e. its likely contribution to the cause of death) was classified as significant, equivocal, or incidental, based on review of the incident history and the pathological, microbiological and parasitological findings for those examined post mortem (see Supplementary Methods for full definitions). For the two captive zoo birds with *E. albertii* isolated from clinical samples, the history and ancillary diagnostic test results were also appraised to infer likely isolate significance to host health.

Genome sequencing and quality control

Isolates of *E. albertii* from UKHSA and ZSL were sequenced at GBRU according to previously described protocols (28) and deposited in the Sequence Read Archive (SRA) under the bioproject accession PRJNA315192 with the SRA accession numbers of individual isolates listed in Supplementary Table 1. Short read sequences were retrieved from the SRA and processed using Trimmomatic v0.38 (30) to trim adaptors and filter low-quality bases. FastQC v0.11.6 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and MultiQC v1.7 (31) were used to assess the quality of reads.

Phylogenetic and clustering analysis

Processed reads were mapped to the *E. albertii* strain 1551-2 reference genome (GenBank accession CP025317) (32) using BWA mem v0.7.17 (33). Alignment files were sorted and filtered using the SAMtools suite v1.9-47 (34), PCR duplicates were marked using Picard v2.21.1-SNAPSHOT MarkDuplicates (http://broadinstitute.github.io/picard/). The BCFtools suite v1.9-80 (34) was used to identify sequence variants and filter variant files, in which low quality single nucleotide polymorphisms (SNPs) were removed if mapping quality < 60, Phred-scaled quality score < 30, read depth < 10 and variant allele frequency < 0.7.

BCFtools consensus was used to generate reference-based pseudogenomes for each isolate from the filtered SNP variants. Regions containing insertion sequences and phages (identified using the PHASTER web server https://phaster.ca/) were identified from the reference genome and masked using BEDTools v2.28.0 maskfasta (35). Regions with read depth of < 10 were also masked. The masked pseudogenomes were concatenated and provided as an alignment for Gubbins v2.3.4 (36) to identify and mask regions of putative recombination. Following Gubbins, SNP-sites v2.4.1 (37) was used to extract variant sites, producing a final SNP-alignment of 26,594 bp in length. This SNP-alignment was used to construct a maximum-likelihood phylogenetic tree using IQ-TREE v2.0-rc2 (38), constructed based on the FreeRate nucleotide substitution, invariable site, and ascertainment bias correction model with 1000 bootstrap replicates. The phylogenetic tree was midpoint rooted and visualised using interactive Tree of Life (iTOL) v6.5 (39).

RhierBAPS v1.1.3 (40) was used to identify clusters of genetically similar isolates among the SNP-alignment, termed Bayesian Analysis of Population Structure (BAPS) clusters.

Construction of cgMLST tree with publicly available data

To deepen the insights gained from the UKHSA and ZSL E. albertii isolates, we analysed their genome sequences in the context of publicly available E. albertii sequence data. Specifically, additional publicly available E. albertii genome sequences accessible through Enterobase on the 7th of February 2022 (n = 475) were constructed alongside the data above into a core genome Multi Locus Sequence Type (cgMLST) tree using hierarchical clustering (HeirCC) (41). Minimal metadata on source and country of origin was extracted from Enterobase alongside HeirCC level classifications and visualised over the unrooted cgMLST tree using interactive Tree of Life (iTOL) v6.5 (39). Metadata on isolate origin was manually curated into the following categories: human, avian (poultry, non-poultry and not defined); mammal (livestock, wildlife and companion species); food, water and undescribed sources.

AMR and virulence gene analysis

Draft genomes were assembled de novo from processed short read sequences using Unicycler v0.4.7 (42) with –min_fasta_length set to 200. Qualities of the draft assemblies were assessed with QUAST v5.0.2 (43) and were all within the assembly quality standards of EnteroBase for *Escherichia* (41). Prokka v1.13.3 (44) was used to annotate draft genome sequences.

The presence of known genetic determinants of AMR was detected using AMRFinderPlus v3.9.3 (45) and screened against the Pathogen Detection Reference Gene Catalog (https://www.ncbi.nlm.nih.gov/pathogens/). AMRFinderPlus was run with the organism-specific option for *Escherichia* and screening for both point mutations and genes (with 80% coverage and 90% identity threshold applied). AMR resistance profiles were visualised with UpSetR v2.1.3 (46).

Association of known AMR genes with related plasmid sequence were identified by extracting AMR-gene containing contiguous sequences from draft genome assemblies and comparison against the NCBI nonredundant database using MegaBlast.

Detection of virulence genes was performed using ABRicate (https://github.com/tseemann/abricate), by which draft genomes were screened against the Virulence Factor Database with minimum nucleotide identity of 60% and minimum coverage of 60%. This screen comprised of virulence genes associated with *E. albertii* including *stx*, *eae* and *cdtABC* genes that encode Shiga toxin, intimin and CDT.

Statistical testing

Statistical support for phylogenetic clustering of bird and human isolates was evaluated with chi-square testing on: 1] the proportion of human isolates in individual clusters (Table 1) and 2] associations of HACs and BACs with patient age (categorised into infant [< 2 yrs], children [2–15 yrs], adult [16–60 yrs] and elderly [60 > yrs]). In the finch (Fringillidae) hosts, any significance between the presence of *stx*2f and clinically significant *E. albertii* infection was also evaluated using the Fisher's exact test. Adjusted and strata-specific odds ratios for the effect of bird family on the association between *stx2f* presence and inferred significant disease was conducted using the Mantel-Haenszel Test. All statistical tests were performed using R v4.0.3.

Table 1
Phylogenetic and epidemiological features of *Escherichia albertii* BAPS clusters

	Genomic features		Isolate composition				Statistical support and nomenclature	
BAPS cluster	Congruence with phylogeny	Average Pairwise distance	Total isolates (n)	Human (% of cluster)	Wild bird (% of cluster)	Captive bird (% of cluster)	Proportion humans [Proportion (95% CI), two tailed p-value]	Final determination^
1	Monophyletic	79	4	4 (100)	0 (0)	0 (0)	1.00 (0.40–1.00), p = 0.1232	HAC
2	Monophyletic	28	17	17 (100)	0 (0)	0 (0)	1.00 (0.80–1.00), p < 0.0001	HAC
3	Monophyletic	132	7	7 (100)	0 (0)	0 (0)	1.00 (0.59–1.00), p = 0.0156	HAC
4	Monophyletic	25	16	16 (100)	0 (0)	0 (0)	1.00 (0.79–1.00), p < 0.0001	HAC
5	Monophyletic	167	4	4 (100)	0 (0)	0 (0)	1.00 (0.40–1.00), p = 0.1232	HAC
6	Monophyletic	967	61	12 (20)	49 (80)	0 (0)	0.20 (0.11–0.33), p < 0.0001	BAC
7	Monophyletic	73	29	4 (14)	24 (86)	1 (3)	0.17 (0.05–0.35), p = 0.0003	BAC
8	Polyphyletic	3169	24	19 (79)	1 (4)	4 (17)	0.80 (0.58–0.93), p = 0.0148	HAC
Total	NA	NA	162	85	74	5		

^BAC = Bird associated cluster, HAC = human associated cluster

Phenotypic antimicrobial resistance testing

Minimum Inhibitory Concentration (MIC) determination was carried out using Lioflichem® MIC test strips (Liofilchem, Italy) following the manufacturer's guidelines. Bacterial inoculum for MIC testing was prepared, following the EUCAST guidelines for *Enterobacterales* standard broth microdilution (https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_11.0_Breakpoint_Tables.pdf) and was spread on Mueller Hinton Agar plates (Bio-Rad, France) using sterile cotton swabs after which the MIC test strip was applied. Plates were incubated at 37 °C for 18 hours before the readings were recorded.

Results

Summary of the human isolates

Between January 2014 (when routine WGS was first implemented at the GBRU) and December 2021, 83 isolates from human cases were confirmed as *E. albertii*. Over this 8-year period, between 4 and 23 isolates were identified per year (Supplementary Fig. 1). Metadata regarding patient gender, age and history of recent travel were available for 82, 83 and 26 isolates, respectively (Supplementary Table 1). There was no statistical association of isolates with gender (39 males, 32 females) and little association with age group (Table 1/Fig. 1/Supplementary Fig. 1). A total of 24 (29%, $n = 24/83$) patients stated they had recently travelled (within 7 days of onset of symptoms) outside the UK, of which the majority ($n = 21/24$, 88%) reported travel to Asia. Travel status was unknown for the remaining cases (71%, $n = 59/83$), as their travel history was not recorded.

Summary of the bird isolates

Seventy-four *E. albertii* isolates from wild birds were analysed over the period 2000–2019 inclusive. With a single exception (tawny owl *Strix aluco*), the hosts were Passeriformes from the following families in declining rank order: Fringillidae n = 50, Passeridae n = 8, Turdidae n = 7, Paridae n = 4 and single birds from the Hirundinidae, Motacillidae, Prunellidae, and Sturnidae (for species composition see Supplementary Table 2). Isolates were identified each year across the 20-year study period with two exceptions and from a total of 72 sites. Available data permitted determination of the inferred significance of *E. albertii* to host health (see Supplementary methods) for 69 wild birds; with 38% ($n = 26$) being significant, 46% ($n = 32$) being equivocal and 16% ($n = 11$) being incidental. The wild birds for which *E. albertii* infection was considered significant to host health comprised Fringillidae (bullfinch *Pyrrula pyrrhula* $n = 1$, chaffinch *Fringilla coelebs* $n = 4$, greenfinch *Chloris chloris* $n = 9$ and siskin *Spinus spinus* $n = 8$), house sparrow *Passer domesticus* $n = 3$ and a single blue tit *Cyanistes caeruleus*.

The five isolates from captive zoo birds were from a diverse range of species (Anseriformes, Passeriformes, Pelecaniformes, and Sphenisciformes). Inferred significance to host health was categorised as significant for one captive bird (black-footed penguin *Spheniscus demursus*), equivocal for two cases and incidental for two cases.

Fringillidae were more frequently associated with 'significant' inferred clinical significance than non-Fringillidae species combined ($p = 0.0433$, Fisher's exact test) for the wild and captive bird data, and this was also well supported statistically among wild birds alone (p = 0.0612).

Genomic epidemiology of Escherichia albertii from humans and bird isolates

To explore the genomic epidemiology of *E. albertii* among the human and bird isolates from GB, demographic features were overlayed on the bacterial population structure and statistical support for associations with metadata variables were evaluated.

Specifically, to determine the population structure of *E. albertii* within our dataset, a maximum likelihood phylogeny was constructed based on a SNP alignment of 26,594 bp (Fig. 1). BAPS identified eight clusters consistent with monophyletic clustering, with the exception of BAPS cluster 8, which was split across multiple regions of the tree (Table 1, Fig. 1). Combining the epidemiological information with this population structure revealed distinct and separate phylogenetic clustering of bird and human isolates ($p < 0.0001$, Chi-square test, 7 df), although statistical support varied for individual clusters (see Table 1). Most bird isolates ($n = 74/79$) belonged to BAPS clusters 6 and 7 in which bird isolates were statistically over-represented, and these were termed bird-associated clusters (BACs, Table 1). To facilitate further high-level investigation, BAPS clusters 1,2,3,4,5 and 8 were termed human-associated clusters (HACs). Intermixing between human and bird isolates was observed within both BACs and one HAC. Specifically, the HAC BAPS 8 contained 6% ($n = 5/79$) of isolates from birds, 4/5 of which were from captive zoo birds. Within the BACs 6 and 7, 18% ($n = 16/90$) of isolates were from humans.

To investigate the association of *E. albertii* with human demographic features, we associated travel history and patient age with the bacterial population structure. All 24 isolates from human patients with a confirmed recent history of international travel belonged to HACs, and at least one travel-associated isolate was identified in each of the six HACs (Fig. 1). The travel status was not recorded for any of the human cases with isolates that fell within the BACs. When associating human age groups with population cluster assignation (BAC/HAC), we observed a significant difference between the BACs and HACs ($p = 0.0008$, Fisher's exact, Fig. 1, Supplementary Table 3). Within the BACs, infant (< 2 years) and elderly (60 > years) were the predominant human age groups, comprising 44% ($n = 7/16$) and 31% ($n = 5/16$) of human isolates respectively (where patient age information was available, Supplementary Table 1). In contrast, the predominant age group within the HACs was adult (16–60 years) comprising 55% ($n = 37/67$) of human isolates.

Virulence profiles and associations with disease in bird hosts

The *eae* gene was present in all but one isolate within the dataset, and the *cdtA*, *cdtB*, *cdtC* genes were present in > 94% ($n = 153/162$) isolates (Fig. 2). The *stx2f* gene was detected in 38 isolates, the majority ($n = 37/38$, 97%) of which were from wild birds in BACs, except for one human isolate (SRR6144114) belonging in BAPS 8. Among the wild birds, *stx2f* resulted in an increased odds of inferred clinical significance of infection (relative to equivocal and incidental combined) (OR = 10.27, 95% CI = 2.98–35.45 $p = 0.0002$). There was little evidence for confounding of the disease association by bird family (Fringillidae/Non-Fringillidae, Adjusted OR 10.25 95% CI 2.66–92.78), a possible effect modification of the bird family (Strata specific OR: OR = 12.68, 95% CI = 2.66–877.38, p-value < 0.001 (Fringillidae), OR = 0.64 95% CI 0.03–16.03, p-value = 1). This was challenging to evaluate further as the *stx2f* was over-represented among the Fringillidae (vs non-Fringillidae OR = 25.67, 95% CI = 5.35-123.23. $p = 0.0001$), specifically of 37 *stx2f*-positive bird isolates, 35 were from Fringillidae species.

Antimicrobial resistance profiles in human and bird isolates

To investigate the genotypic predictors of AMR among *E. albertii* isolates in this dataset, we looked for the presence of genetic determinants of AMR. Both horizontally acquired antimicrobial resistance genes (ARGs) and vertically inherited point mutations known to confer resistance or reduced susceptibility to various antimicrobials in *E. coli* were identified. ARGs were exclusively identified in human isolates, except for one captive zoo bird isolate in HAC BAPS cluster 8 (SRR13092475). Overall, human isolates were observed to carry more AMR genetic determinants compared to bird isolates, including a total of 25 ARGs and five point mutations associated with resistance or reduced susceptibility to 10 different antimicrobial drug classes. In contrast, only three point mutations were identified among the bird isolates, with the exception of the aforementioned captive bird isolate (SRR13092475) carrying an additional 6 ARGs associated with resistance to 7 antimicrobial drug classes. Point mutations were more frequent than ARGs, but the implications less clear. Specifically, *uhpT* E350Q and I355T (Fig. 3a), predicted to confer resistance against fosfomycin and quinolone, respectively (47), were identified in all human and bird isolates, with the multidrug-resistance associated *marR* S3N point mutation being identified in the majority ($n = 157/162$, 97%) of isolates.

The genotypic AMR profile among human isolates was further explored through phenotypic testing. There were 18 unique genotype profiles, including three dominant profiles identified in 80% ($n = 66/83$) of the isolates (Fig. 3b). Correlating ARGs with the phylogeny revealed that the majority (14/16) of isolates within the HAC BAPS 4 had the ARGs *blaDHA-1*, *blaTEM-1*, *dfrA17*, *mph(A)*, *qnrB4*, *sul1* and *tet(A)* (Fig. 1). Among these, *mph(A)*, *sul1*, *blaDHA-1* and *qnrB4* were present on a single contig in multiple isolates, the longest of which was 14,961 bp. A BLASTn search of this contiguous sequence revealed 100% coverage and identity with plasmids from multiple *E. coli* strains, *Shigella sonnei* and *S. flexneri* (Supplementary Table 3). Single contiguous sequences

containing the 4 ARGs were identified in 13 isolates, all belonging to BAPS cluster 4. A single point mutation in the quinolone resistance determining region (QRDR) of gyrA, S83L, was present in 60% (43/72) of HACs isolates (though this was not present in BAPS 3) and only 1% (1/90) of isolates in BACs (Fig. 1).

Antimicrobial resistance profiles of 11 *E. albertii* (HAC *n* = 7, BAC *n* = 4) isolates for cefoxitin, ceftriaxone, fosfomycin, tetracycline, ampicillin, ciprofloxacin, chloramphenicol and rifampicin were tested to review the phenotypic consequences of mutations identified in this study (Figs. 1 & 3). The presence of ARGs *tet(A)*, *blaTEM-1* and *qnrB4* conferred resistance to tetracycline, β-lactam and fluoroquinolone class antibiotics respectively (Table 2). The presence of ARG *blaDHA-1* did not confer resistance to the cephalosporin class antibiotics, cefoxitin and ceftriaxone, in this isolate set. Point mutations *uhpA_G97D** and *uhpT_E350Q**, when present together, as well as point mutations in *gyrA_S83L**, were associated with resistance/decreased susceptibility to their related antimicrobial classes (fosfomycin and fluroquinolones respectively). Point mutations in *marR_S3N** and *parE_I355T** were present in the majority of isolates tested in this study set, and resistance profiles were consistent across the dataset and impacted by the additional presence of other ARGs or point mutations (Table 2)*Global contextualisation of E. albertii from GB*

Table 2
Antimicrobial resistance phenotypes of 11 *E. albertii* isolates.

SRA Accession	Genotype [a]	MIC μgmL^{-1} [c]							
		Cephalosporin		Fosfomycin	Tetracycline	β-lactam	Fluoroquinolone	Chloramphenicol	Rifan
		Cefoxitin	Ceftriaxone	Fosfomycin	Tetracycline	Ampicillin	Ciprofloxacin	Chloramphenicol	Rifan
SRR12769799	*uhpT_E350Q*, marR_S3N*, parE_I355T**	8	0.047	1	1.5	3	0.016	6	8
SRR12769953	*uhpT_E350Q*, marR_S3N*, parE_I355T**	6	0.047	6	0.5	4	0.008	3	2
SRR13049225	*uhpT_E350Q*, marR_S3N*, parE_I355T**	6	0.047	2	1	4	0.006	4	4
SRR13049237	*uhpT_E350Q*, parE_I355T**	6	0.047	1	1	4	0.012	6	3
SRR11442290	*uhpT_E350Q*, marR_S3N*, parE_I355T*, gyrA_S83L**	4	0.047	1.5	0.75	6	0.125	3	12
SRR15338008	*uhpT_E350Q*, marR_S3N*, gyrA_S83L*, parE_I355T**	1.5	< 0.016	1.5	0.38	3	0.032	2	4
SRR8981835	*uhpT_E350Q*, marR_S3N*, gyrA_S83L*, parE_I355T**	1.5	< 0.016	1.5	0.25	3	0.032	2	4
SRR15338057	*blaDHA-1, uhpA_G97D*, uhpT_E350Q*, tet(A), marR_S3N*, blaTEM-1, gyrA_S83L*, parE_I355T*, qnrB4*	6	0.064	12	48	> 256	0.38	1.5	4
SRR9050433	*blaDHA-1, uhpA_G97D*, uhpT_E350Q*, tet(A), marR_S3N*, blaTEM-1, gyrA_S83L*, parE_I355T*, qnrB4*	1	0.047	8	48	96	0.5	3	16
SRR11425059	*uhpT_E350Q*, tet(A), marR_S3N*, blaTEM-1, gyrA_S83L*, parE_I355T**	12	0.064	2	48	> 256	0.19	4	8
SRR3574322	*uhpT_E350Q*, tet(A), marR_S3N*, blaTEM-1, gyrA_S83L*, parE_I355T**	3	< 0.016	1	32	96	0.023	4	4

[a] Genes and point mutations (*) found present in isolates, in this study

[b] Genes and point mutations (*) associated with resistance to antimicrobial classes tested in this study

[c] Underlined MIC determination results highlight MIC breakpoints (μgmL^{-1}) classed as resistant according to EUCAST guidelines (https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_11.0_Breakpoint_Tables.pdf)

Supplementary Fig. 1. Number of human and bird *Escherichia albertii* isolates per year. Human isolates were collected 2015–2021 (blue), while bird isolates w collected 2000–2019 inclusive (pink).

Supplementary Table 3. Results from BLASTn search of contiguous sequence identified in 14 *Escherichia albertii* isolates within a monophyletic clade. The ta lists plasmids that share 100% coverage and sequence identity with the 14,961 bp sequence from isolate SRR8838300 harbouring *mph(A)*, *sul1*, *blaDHA-1* an *qnrB4*.

SRA Accession	Genotype [a]	MIC µgmL^{-1} [c]							
		Cephalosporin		Fosfomycin	Tetracycline	β-lactam	Fluoroquinolone	Chloramphenicol	Rifan
		Cefoxitin	Ceftriaxone	Fosfomycin	Tetracycline	Ampicillin	Ciprofloxacin	Chloramphenicol	Rifan
Genotype associated with resistance [b]		blaDHA-1		uhpA_G97D*, uhpT_E350Q*	tet(A), tet(B), marR_S3N*	blaTEM-1, blaTEM-135, marR_S3N*	gyrA_S83L*, parC_S57T*, parE_I355T*, qnrB19, qnrB4, qnrS13, marR_S3N*	marR_S3N*	marR

[a] Genes and point mutations (*) found present in isolates, in this study

[b] Genes and point mutations (*) associated with resistance to antimicrobial classes tested in this study

[c] Underlined MIC determination results highlight MIC breakpoints (µgmL^{-1}) classed as resistant according to EUCAST guidelines (https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_11.0_Breakpoint_Tables.pdf)

Supplementary Fig. 1. Number of human and bird *Escherichia albertii* isolates per year. Human isolates were collected 2015–2021 (blue), while bird isolates v collected 2000–2019 inclusive (pink).

Supplementary Table 3. Results from BLASTn search of contiguous sequence identified in 14 *Escherichia albertii* isolates within a monophyletic clade. The ta lists plasmids that share 100% coverage and sequence identity with the 14,961 bp sequence from isolate SRR8838300 harbouring *mph(A)*, *sul1*, *blaDHA-1* an *qnrB4*.

To place the human and bird *E. albertii* isolates from GB within the global context, we expanded the analysis to include additional isolates retrieved from publicly available data (n = 475, Methods, Supplementary Table 4). A cgMLST tree was generated based on hierarchical clustering of 2513 gene loci. These additional isolates were derived from diverse sources (22% human; 47% Avian ['poultry', 'non-poultry' and 'not defined']; 7% mammal (e.g. livestock, wild species and companion species); 1% food; 2% water and 21% undescribed sources) and locations (18% Americas, 16% Europe, 41% Asia, 1% Africa, 1% Oceania, and remaining 23% unknown). We correlated the position of BAPS clustered isolates from the current study in this broader context and we observed that our isolates were dispersed across most parts of the cgMLST tree, indicating that the GB isolates capture much of the known diversity of *E. albertii*. The cgMLST tree also revealed that while isolates belonging to the HACs BAPS 2, 4 and 5 remained largely within individual clades of the tree alongside other human-derived isolates (Fig. 4), isolates from HAC BAPS 3 clustered with poultry-derived isolates from Asia and the USA. The majority of isolates from the wild BAC BAPS 7 were similarly embedded within a cluster, this time dominated by poultry isolates from Asia. However, isolates from BAC BAPS 6 and HAC BAPS 8 appeared in multiple clades intermixed with isolates derived from various sources. This is consistent with their greater phylogenetic distance relative to other BAPS clusters (Table 1) and suggests that the association of these two BAPS clusters as bird- and human-associated may be less clear.

Discussion

Currently, the notification of cases of GI disease caused by *E. albertii* in GB in both humans and animals is low compared to other well-established pathogens, such as *Campylobacter* and *Salmonella* species (Supplementary Fig. 1) (48, 49). However, it is likely that the number of *E. albertii* diagnoses will increase in line with improvements in molecular diagnostics and the wider adoption of PCR and WGS as tools for the surveillance of GI pathogens. Thus, analysing the data we have now to understand the potential burden to public health, clinical significance, and risk factors in both human and animal hosts is necessary to guide future research and surveillance strategies.

Although enhanced surveillance questionnaires are not conducted for *E. albertii*, the patterns we observed in our genomic epidemiological analyses suggest that *E. albertii* infection in people most likely has similar transmission routes and risk factors to other GI pathogens. Specifically, the proportion of reported travel-association (31%, 24/83) is consistent with those observed for other travel-associated *Enterobactericeae*, including *Shigella* (19–50% for the years 2005–2014) (50) and *Salmonella* (19–32% for the years 2005–2014)(51). Like *Salmonella* and *Campylobacter* species, zoonotic infection may play a role in disease transmission (52). In exploring the possibility of zoonotic infection, we observed that human and bird isolates from GB typically belong to host-associated monophyletic groups which does not support substantial cross-species transmission (i.e. zoonotic or anthroponotic) between wild and zoo birds and humans. The distinction of *E. albertii* strains among these two host groups was further supported by distinct and convergent features of the accessory genome. Specifically, with HACs containing or acquiring ARGs and the occurrence of *stxf2* in BACs. However, it is possible that the acquisition of ARGs in HACs of *E. albertii* may have been confounded by geography as many of the HAC isolates were from patients who had recently travelled to Asia, a known risk factor for the acquisition of ARGs among enteric pathogens, and where convergent evolution of QRDR mutations in enteric pathogens is known to be occurring (53, 54).

However, limited occurrence of human isolates in BACs, and vice versa, prompted further investigation of the potential for zoonotic and anthroponotic infections of *E. albertii*. Regarding evidence of zoonosis, human isolates were found among BACs in the GB data (Table 1). Although some human isolates grouped in BACs BAPS 6 (n = 13), indicating a potential zoonotic source, the sequence data did not support direct transmission from birds to humans within BAPS 6, as the isolates were not phylogenetically closely related (Table 1, Fig. 3). Contrastingly, the human isolates that grouped in BAPS 7 (n = 4) had comparatively lower diversity and higher similarity with avian isolates, indicating a more-direct association with avian hosts (Fig. 3). Supplementary feeding of garden birds, which typically include finches, is a common pastime in GB that results in a close human-wildlife interface (55) and a zoonotic infection risk has been identified for other bacterial pathogens of wild birds (56, 57). The humans infected with BAC *E. albertii* showed an age bias toward the very young and elderly, consistent with bias towards infant infection previously described for wildlife-associated biotypes of *Salmonella* Typhimurium and *S.* Enteritidis (57, 58). Hence, the four human isolates in BAPS 7 conceivably represent zoonotic infections and good hygiene measures when feeding garden birds are

recommended as routine best practice to help safeguard public health (e.g. wearing gloves and hand washing after handling bird feeders or cleaning bird tables and avoiding direct contact with sick or dead wild birds).

The possibility of anthroponosis was indicated by four of five captive zoo bird isolates clustering within the HAC BAPS 8 rather than the BACs. However, similar to the BAPS 6 human/bird mixing above, the sequence data did not provide evidence of direct transmission, as the genomic divergence between isolates in BAPS 8 was high (Table 1) and there were other potential sources of *E. albertii* infection among the captive birds. Specifically, these four captive zoo birds were kept in enclosures with outdoor access and various diets (i.e. omnivorous, piscivorous and carnivorous) from a range of providers so had exposure to sympatric captive species, free-living wildlife, and various potential dietary or environmental sources. Ultimately, the limited/imperfect sampling in our study and unknown/unharmonized provenance of the public data precludes firm conclusions on the risk of cross-species transmission of *E. albertii*, but suggests the link is worthy of continued investigation.

Recent studies from China, Japan and the USA have highlighted the potential for foodborne transmission of *E. albertii* to humans, via the consumption of poultry (8, 23, 25). In our study, one of the HACs (BAPS 3) was admixed with poultry-associated isolates from China and the USA (Fig. 3), indicating the possibility that both travel-associated and some domestically acquired *E. albertii* infections may be a foodborne illness linked to eating poultry. BAC BAPS 7 (which contained isolates from humans) was similarly encompassed in a broader group of poultry isolates. In contrast to BAPS 6 and BAPS 7 being embedded among poultry strains, HACs BAPS 2 and BAPS 4 were on long branches without close associations with other hosts or regions (Fig. 3). This, combined with the sporadic nature of *E. albertii* infection in humans, likely indicates an unobserved reservoir of disease, either overseas, or in domestic non-human hosts. If domestic transmission of *E. albertii* from poultry to humans were occurring in Great Britain, we might plausibly have expected GB poultry isolates to populate near the BAPS 2 and BAPS 4 clusters. The emerging picture of *E. albertii*, a travel-associated pathogen with a primary reservoir in poultry, would parallel other enteric pathogens such as *Salmonella* and *Campylobacter* (48, 49). Genomic surveillance of *E. albertii* in a greater number of locations and non-human potential reservoirs is needed to further elucidate the ecology of this pathogen.

Our results and the global publicly available data, combined with the published literature on *E. albertii* (2, 21, 23, 59), indicate that avian hosts play a larger role in the epidemiology of *E. albertii* in human beings than do other (e.g. mammalian) hosts. The metadata of publicly available isolates revealed that comparatively few isolates were derived from mammals relative to avian (7% vs 47% respectively). Although public data are not a reflection of representative surveillance, additional data from the ZSL provide a similar picture. Specifically, the same microbiological protocol has been used for all samples submitted from the diverse taxonomic range of birds and mammals held in the zoological collections at ZSL since 1991, both from clinical samples and routine health checks (ZSL, unpublished data). That *E. albertii* was recovered from only five captive birds and not from any mammal supports a skew of this bacterium towards avian hosts. Furthermore, there have been no confirmed detections of *E. albertii* infection recovered from livestock or wildlife species disease surveillance conducted by the Animal Plant & Health Agency in England and Wales for at least the past 23 years (APHA, unpublished data). The adoption of molecular methods for *E. albertii* surveillance in veterinary laboratories is recommended to further aid understanding of the host species range and the clinical significance of this bacterium in animal hosts, particularly in avian species.

With regard to the implications of *E. albertii* for bird health, our study found that infection was more frequently associated with significant disease in finch than non-finch species. This is consistent with historical investigations of multiple mortality incidents of finches in Scotland and the USA (2, 19) and supports the hypothesis that it acts as a primary pathogen in these wild birds. This bird family bias may be related to host and environmental factors such as differential exposure or susceptibility, and our data indicate that differences in virulence determinant components among the circulating strains of *E. albertii* may also play a role. Specifically, isolates containing the virulence factors *stx2f* were associated with finch hosts (Fringillidae), and infection in these birds was significantly more likely to be associated with disease. This could not be further untangled here owing to the low occurrence of stx2f strains from non-finch species, but it is possible that finches may act as a reservoir of *Stx2f*-positive *E. albertii* in GB, as is hypothesised to occur with garden bird-associated biotypes of *Salmonella* Typhimurium in GB (29). The broader study from which the finch isolates were derived examined ~ 4000 wild bird carcasses, (14 orders of which ~ 85% were Passeriformes) over the period 2000–2019 inclusive across GB (ZSL, unpublished data). This revealed that *E. albertii* infection occurred much less frequently than diseases such as trichomonosis and salmonellosis (60), with the observed pattern being consistent with sporadic endemic disease in finches with a low frequency of occurrence. However, since *E. albertii* infection was detected in a range of wild and captive bird species where inferred significance to host health was incidental or equivocal in this study, and from a range of bird families (e.g. Anatidae, Corvidae, Laridae) in the available global dataset, further surveillance is required to identify the role that different bird families play in the ecology of the bacterium.

In conclusion, due to poor molecular diagnostic capabilities for *E. albertii* in both human and animal health laboratories, the true public health burden of *E. albertii* infection is likely to be underestimated. Furthermore, the lack of systematic surveillance data means that clinical severity and exposure risks are largely unknown. However, we leveraged available data to highlight the likely relevance of travel to regions with a high risk of GI infections, including an association with AMR, and a potential zoonotic component that is likely bird associated, but probably more so with poultry than with wild bird species. To improve surveillance for *E. albertii*, we recommend increased deployment of molecular methods in medical and veterinary diagnostic laboratories, in conjunction with the systematic collection of epidemiological data. Maintaining close collaborations between medical and veterinary institutions, and the integration of human and animal surveillance datasets, is essential to better understand the source, transmission and risks to animal and public health of this recently identified pathogen.

Declarations

Acknowledgements

We thank the members of the public and participants in the British Trust for Ornithology's Garden BirdWatch who assisted with reporting observations of wild bird mortality to ZSL as part of a national wild bird disease surveillance programme, called the Garden Bird Health *initiative* (2005-2012) which was then

integrated into the Garden Wildlife Health project (2013-2019) (www.gardenwildlifehealth.org). We thank the veterinarians who conducted some of the wild bird post-mortem examinations included within this study (including Katie Beckmann, Lydia Franklinos, Joseph Heaver and Vicky Wilkinson) and Ed Fullick from the Animal & Plant Health Agency for assistance with information sharing.

Funding

Financial support for the Garden Bird Health *initiative* (2005-2012) came from the Birdcare Standards Association, British Trust for Ornithology, British Veterinary Association Animal Welfare Foundation, CJ Wildbird Foods, Cranswick Pet Products, UK Department for the Environment Food & Rural Affairs (Defra) and Welsh Government through the Animal & Plant Health Agency's (APHA) Diseases of Wildlife Scheme (DoWS) Scanning Surveillance Programme (Project ED1058), Gardman Ltd, Institute of Zoology, Royal Society for the Protection of Birds and the Universities Federation for Animal Welfare.

Financial support for the Garden Wildlife Health Project (2013-2019 inclusive) came in part from the Department for Environment, Food and Rural Affairs (Defra), the Welsh Government and the Animal and Plant Health Agency (APHA) Diseases of Wildlife Scheme (DoWS); and from the Banister Charitable Trust, Esmée Fairbairn Foundation, Garfield Weston Foundation and the Universities Federation for Animal Welfare. IoZ staff receive financial support from Research England.

This work was supported by a Medical Research Council Grant (KB: MR/R020787/1) and Biotechnology and Biological Resources Council Grant (KB: BB/V009184/1). Kate Baker, David Greig and Claire Jenkins are affiliated to the National Institute for Health and Care Research Health Protection Research Unit (NIHR HPRU) in Gastrointestinal Infections at University of Liverpool in partnership with the United Kingdom Health Securities Agency (UKHSA), and the University of Warwick. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care, or UKHSA.

Ethics statement

Samples were collected during post-mortem examination of wild birds found dead or euthanased for welfare reasons under the Veterinary Surgeons Act 1981.

Data availability

Phylogenetic trees from this study have been deposited in FigShare (DOI: 10.6084/m9.figshare.20894854).

References

1. Huys G, Cnockaert M, Janda JM, Swings J. Escherichia albertii sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. Int J Syst Evol Microbiol. 2003;53(Pt 3):807-10.

2. Oaks JL, Besser TE, Walk ST, Gordon DM, Beckmen KB, Burek KA, et al. Escherichia albertii in wild and domestic birds. Emerging infectious diseases. 2010;16(4):638.

3. Ooka T, Seto K, Kawano K, Kobayashi H, Etoh Y, Ichihara S, et al. Clinical significance of Escherichia albertii. Emerg Infect Dis. 2012;18(3):488-92.

4. Hyma KE, Lacher DW, Nelson AM, Bumbaugh AC, Janda JM, Strockbine NA, et al. Evolutionary genetics of a new pathogenic Escherichia species: Escherichia albertii and related Shigella boydii strains. J Bacteriol. 2005;187(2):619-28.

5. Nimri LF. Escherichia albertii, a newly emerging enteric pathogen with poorly defined properties. Diagn Microbiol Infect Dis. 2013;77(2):91-5.

6. Zhang X, Das S, Dunbar S, Tang YW. Molecular and non-molecular approaches to etiologic diagnosis of gastroenteritis. Adv Clin Chem. 2020;99:49-85.

7. Boxall MD, Day MR, Greig DR, Jenkins C. Antimicrobial resistance profiles of diarrhoeagenic Escherichia coli isolated from travellers returning to the UK, 2015-2017. J Med Microbiol. 2020;69(7):932-43.

8. Luo L, Wang H, Payne MJ, Liang C, Bai L, Zheng H, et al. Comparative genomics of Chinese and international isolates of Escherichia albertii: population structure and evolution of virulence and antimicrobial resistance. Microb Genom. 2021;7(12).

9. Hassan J, Awasthi SP, Hatanaka N, Okuno K, Hoang PH, Nagita A, et al. Development of a multiplex PCR targeting eae, stx and cdt genes in genus Escherichia and detection of a novel cdtB gene in Providencia rustigianii. Pathog Dis. 2018;76(9).

10. Hinenoya A, Ichimura H, Awasthi SP, Yasuda N, Yatsuyanagi J, Yamasaki S. Phenotypic and molecular characterization of Escherichia albertii: Further surrogates to avoid potential laboratory misidentification. Int J Med Microbiol. 2019;309(2):108-15.

11. Slater SL, Sagfors AM, Pollard DJ, Ruano-Gallego D, Frankel G. The Type III Secretion System of Pathogenic Escherichia coli. Curr Top Microbiol Immunol. 2018;416:51-72.

12. Hinenoya A, Yasuda N, Mukaizawa N, Sheikh S, Niwa Y, Awasthi SP, et al. Association of cytolethal distending toxin-II gene-positive Escherichia coli with Escherichia albertii, an emerging enteropathogen. Int J Med Microbiol. 2017;307(8):564-71.

13. Gomes TAT, Ooka T, Hernandes RT, Yamamoto D, Hayashi T. Escherichia albertii Pathogenesis. EcoSal Plus. 2020;9(1).

14. Bhatt S, Egan M, Critelli B, Kouse A, Kalman D, Upreti C. The Evasive Enemy: Insights into the Virulence and Epidemiology of the Emerging Attaching and Effacing Pathogen Escherichia albertii. Infect Immun. 2019;87(1).

15. Ooka T, Tokuoka E, Furukawa M, Nagamura T, Ogura Y, Arisawa K, et al. Human gastroenteritis outbreak associated with Escherichia albertii, Japan. Emerging infectious diseases. 2013;19(1):144.

16. Masuda K, Ooka T, Akita H, Hiratsuka T, Takao S, Fukada M, et al. Epidemiological Aspects of Escherichia albertii Outbreaks in Japan and Genetic Characteristics of the Causative Pathogen. Foodborne Pathog Dis. 2020;17(2):144-50.

17. Vishram B, Jenkins C, Greig DR, Godbole G, Carroll K, Balasegaram S, et al. The emerging importance of Shiga toxin-producing Escherichia coli other than serogroup O157 in England. J Med Microbiol. 2021;70(7).

18. Foster G, Ross HM, Pennycott TW, Hopkins GF, McLaren IM. Isolation of Escherichia coli O86:K61 producing cyto-lethal distending toxin from wild birds of the finch family. Lett Appl Microbiol. 1998;26(6):395-8.

19. Pennycott TW, Ross HM, McLaren IM, Park A, Hopkins GF, Foster G. Causes of death of wild birds of the family Fringillidae in Britain. Vet Rec. 1998;143(6):155-8.

20. Murakami K, Maeda-Mitani E, Kimura H, Honda M, Ikeda T, Sugitani W, et al. Non-biogroup 1 or 2 Strains of the Emerging Zoonotic Pathogen Escherichia albertii, Their Proposed Assignment to Biogroup 3, and Their Commonly Detected Characteristics. Front Microbiol. 2019;10:1543.

21. Hinenoya A, Awasthi SP, Yasuda N, Nagano K, Hassan J, Takehira K, et al. Detection, Isolation, and Molecular Characterization of Escherichia albertii from Wild Birds in West Japan. Jpn J Infect Dis. 2022;75(2):156-63.

22. Oh JY, Kang MS, Hwang HT, An BK, Kwon JH, Kwon YK. Epidemiological investigation of eaeA-positive Escherichia coli and Escherichia albertii strains isolated from healthy wild birds. J Microbiol. 2011;49(5):747-52.

23. Hinenoya A, Li XP, Zeng X, Sahin O, Moxley RA, Logue CM, et al. Isolation and characterization of Escherichia albertii in poultry at the pre-harvest level. Zoonoses Public Health. 2021;68(3):213-25.

24. Sonnevend A, Alali WQ, Mahmoud SA, Ghazawi A, Bharathan G, Melegh S, et al. Molecular Characterization of MCR-1 Producing Enterobacterales Isolated in Poultry Farms in the United Arab Emirates. Antibiotics (Basel). 2022;11(3).

25. Wang H, Zhang L, Cao L, Zeng X, Gillespie B, Lin J. Isolation and characterization of Escherichia albertii originated from the broiler farms in Mississippi and Alabama. Vet Microbiol. 2022;267:109379.

26. Grillova L, Sedlacek I, Pachnikova G, Stankova E, Svec P, Holochova P, et al. Characterization of four Escherichia albertii isolates collected from animals living in Antarctica and Patagonia. J Vet Med Sci. 2018;80(1):138-46.

27. Hinenoya A, Nagano K, Awasthi SP, Hatanaka N, Yamasaki S. Prevalence of Escherichia albertii in Raccoons (Procyon lotor), Japan. Emerg Infect Dis. 2020;26(6):1304-7.

28. Chattaway MA, Schaefer U, Tewolde R, Dallman TJ, Jenkins C. Identification of Escherichia coli and Shigella Species from Whole-Genome Sequences. J Clin Microbiol. 2017;55(2):616-23.

29. Lawson B, Howard T, Kirkwood JK, Macgregor SK, Perkins M, Robinson RA, et al. Epidemiology of salmonellosis in garden birds in England and Wales, 1993 to 2003. Ecohealth. 2010;7(3):294-306.

30. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.

31. Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32(19):3047-8.

32. Romao FT, Hernandes RT, Ooka T, Hayashi T, Sperandio V, Gomes TAT. Complete Genome Sequence of Escherichia albertii Strain 1551-2, a Potential Extracellular and Intracellular Pathogen. Genome Announc. 2018;6(9).

33. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:13033997. 2013.

34. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.

35. Quinlan AR. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. Curr Protoc Bioinformatics. 2014;47:11 2 1-34.

36. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res. 2015;43(3):e15.

37. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microb Genom. 2016;2(4):e000056.

38. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268-74.

39. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47(W1):W256-W9.

40. Tonkin-Hill G, Lees JA, Bentley SD, Frost SDW, Corander J. RhierBAPS: An R implementation of the population clustering algorithm hierBAPS. Wellcome Open Res. 2018;3:93.

41. Zhou Z, Alikhan NF, Mohamed K, Fan Y, Agama Study G, Achtman M. The EnteroBase user's guide, with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and Escherichia core genomic diversity. Genome Res. 2020;30(1):138-52.

42. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017;13(6):e1005595.

43. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072-5.

44. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068-9.

45. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, et al. Validating the AMRFinder Tool and Resistance Gene Database by Using Antimicrobial Resistance Genotype-Phenotype Correlations in a Collection of Isolates. Antimicrob Agents Chemother. 2019;63(11).

46. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics. 2017;33(18):2938-40.

47. Takahata S, Ida T, Hiraishi T, Sakakibara S, Maebashi K, Terada S, et al. Molecular mechanisms of fosfomycin resistance in clinical isolates of Escherichia coli. Int J Antimicrob Agents. 2010;35(4):333-7.

48. LV DEK, Pires SM, Hald T. Attributing foodborne salmonellosis in humans to animal reservoirs in the European Union using a multi-country stochastic model. Epidemiol Infect. 2015;143(6):1175-86.

49. Kaakoush NO, Castano-Rodriguez N, Mitchell HM, Man SM. Global Epidemiology of Campylobacter Infection. Clin Microbiol Rev. 2015;28(3):687-720.

50. Travel-associated Shigella spp. infection in England, Wales and Northern Ireland: 2014. UK Health Security Agency; 2017.

51. Travel-associated non typhoidal Salmonella infection in England, Wales and Northern Ireland: 2014. UK Health Security Agency; 2017.

52. Wang F, Zhang W, Niu D. Editorial: Foodborne Enterobacteriaceae of Animal Origin. Front Cell Infect Microbiol. 2021;11:772359.

53. Sridhar S, Turbett SE, Harris JB, LaRocque RC. Antimicrobial-resistant bacteria in international travelers. Curr Opin Infect Dis. 2021;34(5):423-31.

54. Kantele A, Kuenzli E, Dunn SJ, Dance DAB, Newton PN, Davong V, et al. Dynamics of intestinal multidrug-resistant bacteria colonisation contracted by visitors to a high-endemic setting: a prospective, daily, real-time sampling study. Lancet Microbe. 2021;2(4):e151-e8.

55. Davies ZG, Fuller RA, Loram A, Irvine KN, Sims V, Gaston KJ. A national scale inventory of resource provision for biodiversity within domestic gardens. Biological Conservation. 2009;142(4):761-71.

56. Rehn M, Ringberg H, Runehagen A, Herrmann B, Olsen B, Petersson AC, et al. Unusual increase of psittacosis in southern Sweden linked to wild bird exposure, January to April 2013. Euro Surveill. 2013;18(19):20478.

57. Lawson B, de Pinna E, Horton RA, Macgregor SK, John SK, Chantrey J, et al. Epidemiological evidence that garden birds are a source of human salmonellosis in England and Wales. PLoS One. 2014;9(2):e88968.

58. Lawson B, Franklinos LHV, Rodriguez-Ramos Fernandez J, Wend-Hansen C, Nair S, Macgregor SK, et al. Salmonella Enteritidis ST183: emerging and endemic biotypes affecting western European hedgehogs (Erinaceus europaeus) and people in Great Britain. Sci Rep. 2018;8(1):2449.

59. Murakami K, Etoh Y, Tanaka E, Ichihara S, Horikawa K, Kawano K, et al. Shiga toxin 2f-producing Escherichia albertii from a symptomatic human. Jpn J Infect Dis. 2014;67(3):204-8.

60. Lawson B, Robinson RA, Toms MP, Risely K, MacDonald S, Cunningham AA. Health hazards to wild birds and risk factors associated with anthropogenic food provisioning. Philos Trans R Soc Lond B Biol Sci. 2018;373(1745).

Figures

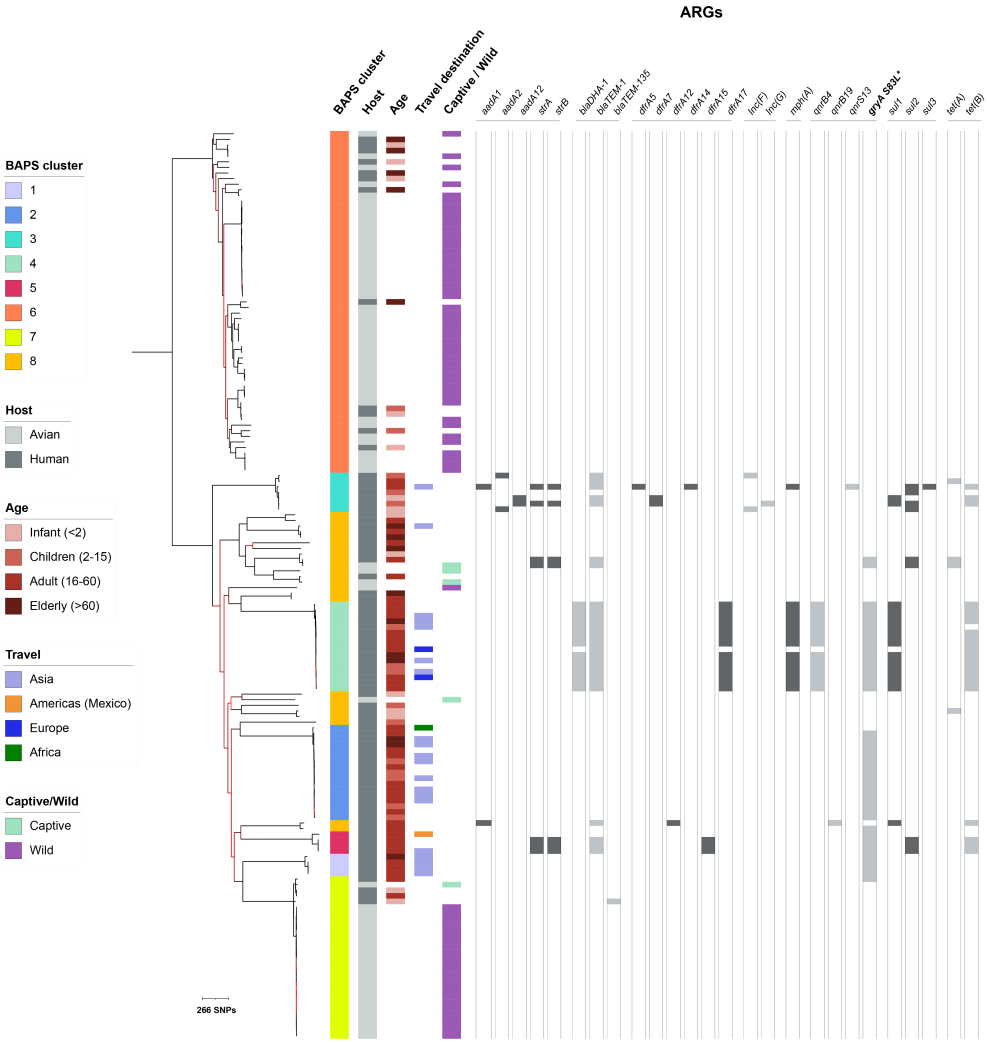

Figure 1

Mid-point rooted maximum likelihood phylogenetic tree of 162 *Escherichia albertii* isolates from Great Britain. Isolate metadata are displayed in the adjacent tracks on the right according to the inlaid keys on the left. Tracks in the centre panel with borders shows presence of antimicrobial resistance genes with the *gryA*S83L point mutation highlighted in bold and indicated with an asterisk. Phylogenetic branches highlighted in red indicates nodes with low bootstrap support between 50 and 70%.

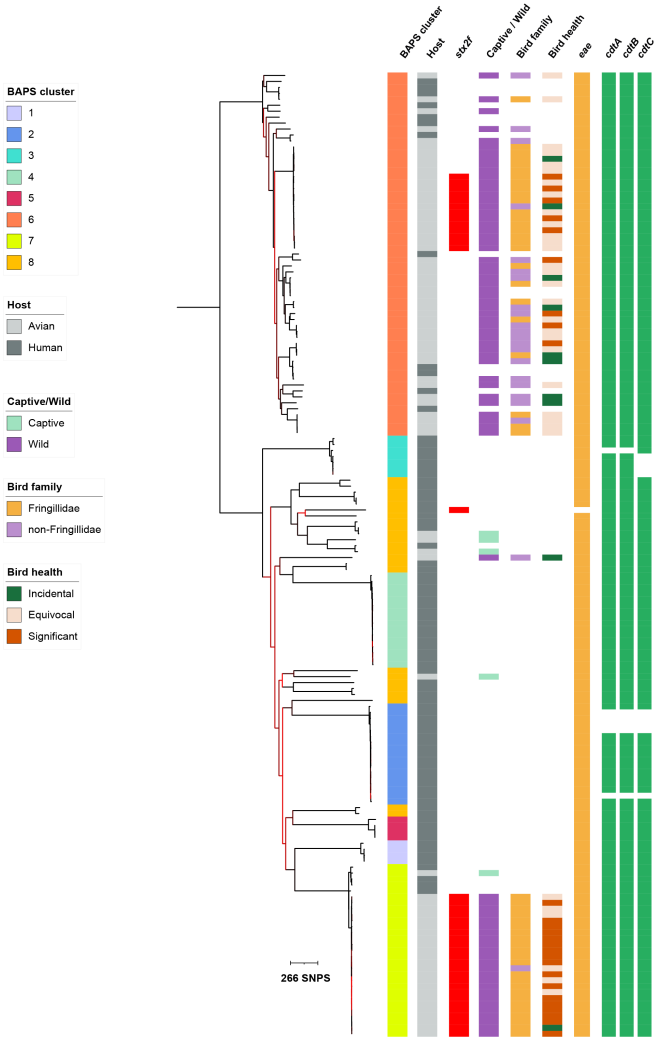

Figure 2

Midpoint-rooted Maximum likelihood phylogenetic tree of Escherichia albertii from Great Britain. Isolate metadata are displayed on the adjacent tracks according to the inlaid key. Phylogenetic branches highlighted in red indicates nodes with low bootstrap support between 50 and 70%.

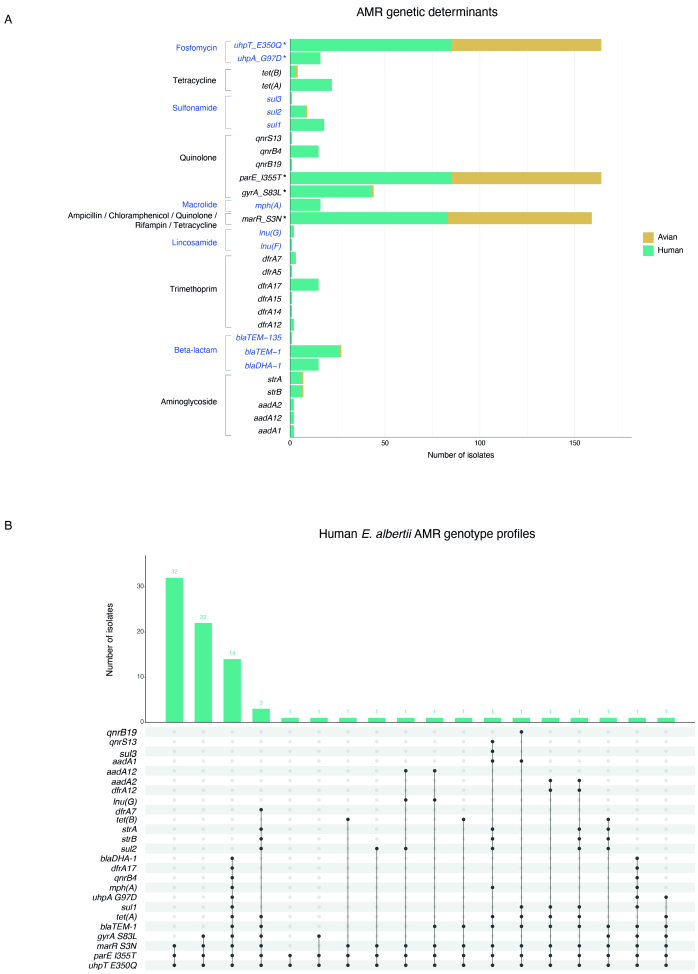

Figure 3

Occurrence of antimicrobial resistance (AMR) among Escherichia albertii isolates from GB. (**A**) Stacked barplot demonstrates the number of isolates from birds and humans carrying known AMR genetic determinants. Genetic determinants highlighted with asterisks represents point mutations and text displayed in alternating colours highlights different antimicrobial drug classes. (**B**) UpSet plot illustrates the prevalence of AMR genotypic profile among human isolates. The combination matrix in the centre panel shows the various genotypic AMR profiles, in which each column represents a unique profile, and each black dot represents presence of a genetic determinant conferring resistance/reduced susceptibility to a drug class (displayed on the left). Vertical barplot above the matrix shows the number of isolates with a particular genotype, the number above each bar shows the exact number of isolates with the genotype.

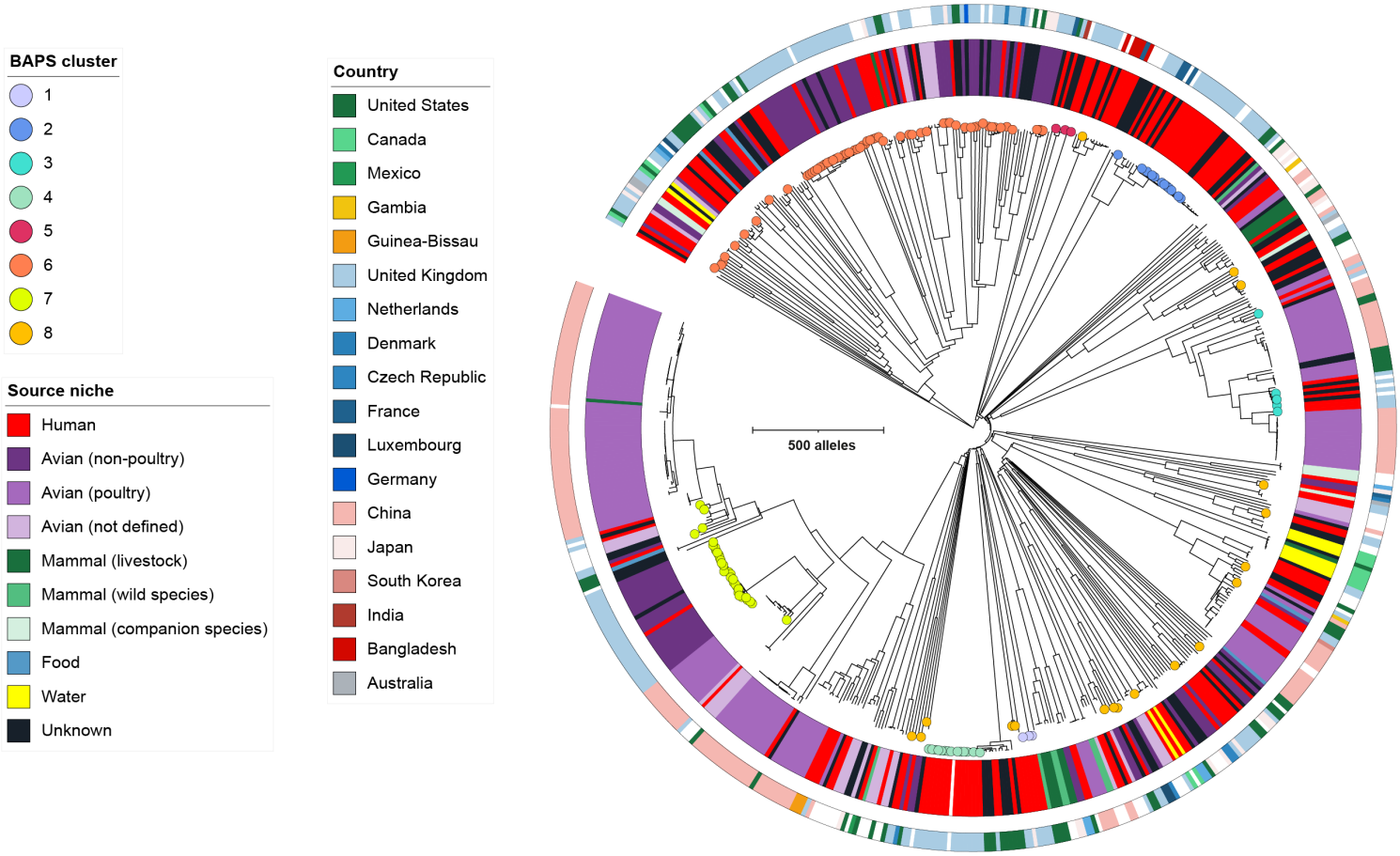

Figure 4

Phylogenetic tree comprising of 162 Escherichia albertii isolates from the current study and an additional 475 isolates retrieved from EnteroBase. Tree was constructed based on MLST sequences. Circles at tree tips highlight E. albertii isolates from Great Britain under investigation in this study, and the colour of the circles represents the BAPS clusters identified earlier in the study. The thicker inner ring, demonstrates the source of the isolates and the thinner outer ring demonstrates the isolate country of origin, all of which are labelled according to the inlaid keys displayed on the left.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigureS1.pdf
- supplementarytablesfinal.xlsx
- SupplementaryTable3.docx