Environment International 186 (2024) 108606

Contents lists available at ScienceDirect

Environment International

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

Full length article

# Changing dynamics of antibiotic resistant *Escherichia* in Caspian gulls shows the importance of longitudinal environmental studies

Kristina Nesporova [a], Michaela Ruzickova [a,b], Hassan Tarabai [a,c], Simon Krejci [b], Martina Masarikova [d], Jarmila Lausova [a,b], Ivan Literak [a,b], Monika Dolejska [a,b,e,f,*]

[a] CEITEC VETUNI Brno, University of Veterinary Sciences Brno, Brno, Czech Republic
[b] Department of Biology and Wildlife Diseases, Faculty of Veterinary Hygiene and Ecology, University of Veterinary Sciences Brno, Brno, Czech Republic
[c] Department of Parasitology, Faculty of Science, University of South Bohemia, Ceske Budejovice, Czech Republic
[d] Department of Infectious Diseases and Microbiology, Faculty of Veterinary Medicine, University of Veterinary Sciences Brno, Brno, Czech Republic
[e] Biomedical Centre, Charles University, Pilsen, Czech Republic
[f] Department of Clinical Microbiology and Immunology, Institute of Laboratory Medicine, The University Hospital Brno, Czech Republic

ARTICLE INFO

Handling Editor: Thanh Nguyen

*Keywords:*
Environmental persistence of AMR bacteria
AMR in wild birds
Successful *Escherichia* STs

ABSTRACT

This study is focused on *Escherichia* spp. isolates resistant to critically important antibiotics (cefotaxime, ciprofloxacin and colistin) among Caspian gull's (*Larus cachinnans*) chicks nesting in the Nove Mlyny Water Reservoir, Czech Republic. The prevalence of antimicrobial resistance (AMR) in bacteria within wild birds is commonly evaluated using a single sampling event, capturing only a brief and momentary snapshot at a particular location. Therefore, the Caspian gulls in our study were sampled in May 2018 (n = 72) and May 2019 (n = 45), and a water sample was taken from the reservoir (2019). We obtained 197 isolates identified as *E. coli* by MALDI-TOF MS. A total of 158 representative isolates were whole-genome sequenced, 17 isolates were then reclassified to *Escherichia albertii*. We observed a higher (86 %; 62/72) occurrence of ESBL/AmpC-producing *Escherichia* spp. among gulls in 2018 compared to 38 % (17/45) in 2019 ($p < 0.00001$). The decrease in prevalence was linked to clonal lineage of *E. coli* ST11893 predominating in 2018 which carried $bla_{CMY-2}$ and which was not recovered from the gulls in 2019. Oppositely, several *Escherichia* STs were found in gulls from both years as well as in the water sample including STs commonly recognized as internationally high-risk lineages such as ST10, ST58, ST88, ST117, ST648 or ST744. Phylogenetic analysis of *E. coli* from EnteroBase from countries where these particular gulls wander revealed that some STs are commonly found in various sources including humans and a portion of them is even closely related (up to 100 SNPs) to our isolates. We demonstrated that the occurrence of AMR in *Escherichia* can vary greatly in time in synanthropic birds and we detected both, a temporary prevalent lineage and several persistent STs. The close relatedness of isolates from gulls and isolates from EnteroBase highlights the need to further evaluate the risk connected to wandering birds.

## 1. Introduction

Contamination of wildlife with antibiotic residues and antibiotic resistant bacteria (ARB) is an increasingly described phenomenon. As wild animals are almost never treated with antibiotics it is evident that the origin of ARB pollution in wildlife was caused by outside factors and it may be ascribed to human activities (Ahlstrom et al., 2021). Any type of anthropogenic pollution in wildlife should be principally seen as an issue, however with the antimicrobial resistance (AMR) the problem is more complex as it may spread via wild animals and the environment and open possibilities for further evolution of the potentially pathogenic bacteria.

An extra focus regarding the ARB should be targeted on wild wandering and migratory birds, especially synanthropic species. Synanthropic birds commonly live in proximity to humans, often feed in dump sites and landfills and interact with diverse water sources (Chytil et al., 2022), therefore they are more prone to be colonised with ARB (Ahmed et al., 2019). This was reported by a previous study from Alaska where they detected a higher level of resistance among gulls coming from urban area compared to the gulls inhabiting an island with limited

* Corresponding author at: University of Veterinary Sciences Brno, Palackeho tr. 1946/1, 612 42 Brno, Czech Republic.
*E-mail address:* monika.dolejska@gmail.com (M. Dolejska).

https://doi.org/10.1016/j.envint.2024.108606
Received 13 December 2023; Received in revised form 21 March 2024; Accepted 25 March 2024
Available online 26 March 2024

anthropogenic influence (Atterby et al., 2016). Similar results were obtained for gulls in Australia (Wyrsch et al., 2022). Additionally, striking genetic similarities were observed in cefotaxime-resistant *E. coli* among gulls and humans in Sweden and in gulls and humans in Australia (Wyrsch et al., 2022; Atterby et al., 2017; Nesporova et al., 2020).

The need to consider AMR as a multi-sectorial and complex issue has been recognised in recent years and framed within the One Health concept (Van Puyvelde et al., 2018). However, data on the environmental dynamics and ecology of AMR bacteria remain limited (Na et al., 2018). Only a few studies address ecological questions, such as the bacteria's ability to live in diverse niches, adapt, and persist over extended periods (Na et al., 2018; Liu et al., 2021). Our attention is still too often narrowed towards antibiotic resistance itself and to clinical settings without searching for the complex picture (Torres et al., 2020).

Caspian gulls count for synanthropic birds and expansive species and they have been extending to Czechia from Ukraine (Chytil et al., 2022). The first breeding pair was observed in 1990, raising up to 248 breeding pairs in 2020, and their first and main colony in the Czech Republic was established in Nove Mlyny Reservoir (Chytil et al., 2022). In some areas including central Europe, Caspian gulls has been declared an invasive species and they exhibit nomadic lifestyle with lengthy dispersal movements independent of season-related migration (Chytil et al., 2022). Therefore, we also described the migratory habits of these gulls within a parallel study which involved telemetry (Chytil et al., 2022). These data were crucial for one of our objectives in this study - to compare data about AMR *Escherichia coli* found in these gulls with publicly available data from countries to which these gulls wander and explore their phylogenetic relations. The main objective was to provide insights into the prevalence dynamics of AMR *E. coli* among Caspian gulls in Nove Mlyny Reservoir in the Czech Republic for two consecutive years. The study aimed to describe the occurrence and persistence of specific *E. coli* sequence types (STs) among gulls and in the water environment of their colony and highlight the factors that may contribute to making certain STs a successfully persistent lineage. We further aimed to describe the mechanisms underlying the AMR and describe mobile genetics elements linked to them.

## 2. Materials and methods

### 2.1. Samples collection and selective cultivation

Cloacal sampling of gulls' nestings was performed on the 24th of May 2018 and 25th of May 2019 from a gulls' colony located on an island in the Nove Mlyny water reservoirs (Chytil et al., 2022). The island with the gulls' colony is located on 48°53′10.6″N 16°36′00.6″E with area of 0.002795 $km^2$. In 2018, the number of breeding pairs was 150, increasing to 211 pairs in 2019 (Chytil et al., 2022). For the time of birds nesting, the visits to the island are restricted so there is limited direct influence of humans on the gull's chicks. A total of 72 cloacal swabs in Amies transport medium (ThermoFisher Scientific, Oxoid, Massachusetts, USA) were collected in 2018 and 45 cloacal swabs were obtained in 2019 from individual gull's chicks (Chytil et al., 2022). In addition, a water sample was taken from the Nove Mlyny water reservoir where the gulls' colony was nesting on 27th July 2019. The samples from gulls were placed into buffered peptone water (ThermoFisher Scientific, Oxoid) and stored at 4 °C for several hours as the primary samples for selective cultivation. The handling of the gulls was in concordance with Ethical standards and covered by approval of the local Czech nature protection authorities (Permissions S-JMK78643/2018 OŽP/Ško and S-JMK 40970/2019 OŽP/Ško).

Cloacal samples from gulls and obtained pellet from centrifuging 1 L of collected water sample were enriched in buffered peptone overnight (10 mL, 37 °C with shaking at 140 RPM). Following that, enriched cultures originating from cloacal samples were cultivated on three plates with MacConkey agar (MCA, Oxoid, UK) and cefotaxime (2 mg/L, Sigma-Aldrich, Missouri, USA), or ciprofloxacin (0.05 mg/L, Sigma-Aldrich), or Eosin methylene blue agar (ThermoFishes Scientific, Oxoid) supplemented with colistin (3.5 mg/L, Sigma-Aldrich). A single colony of preferably *E. coli* morphology was taken from each plate and subsequently inoculated to obtain a pure culture using the same type of media from which the colony originated. The water sample was inoculated on the same type of media as cloacal samples and we aimed to select up to 20 colonies of *E. coli* morphology from each plate for purification. Nevertheless, a sufficient number of colonies was obtained only for cefotaxime-supplemented agar while only a single *E. coli*-like colony was recovered from media with ciprofloxacin. All isolates were species identified using MALDI-TOF MS (Bruker Daltonics, Germany) and only isolates identified as *E. coli* (a total of 197) were included in further testing.

### 2.2. Phenotype characterization and PCR-based detection of selected antimicrobial resistance genes

All *Escherichia* isolates were tested for susceptibility to the set of 18 antimicrobial substances using disc diffusion methods and COLISPOT method as described previously (Sismova et al., 2023). The zones were evaluated with EUCAST (2020) (Committee, 2020) criteria where applicable and the remaining antibiotics were evaluated using CLSI (2017) (Clinical and Laboratory Standards Institute (CLSI), 2500) guidelines. The production of extended-spectrum beta-lactamase (ESBL) and AmpC type beta-lactamase in isolates selected from media with cefotaxime was tested by MAST phenotype diffusion test (D68C1 AmpC & ESBL Detection Set, MAST® Diagnostics, UK). The DNA of isolates from media with ciprofloxacin and colistin was obtained using heat lysis and screened for the presence of transferable mechanism of quinolone resistance (TMQR) genes (*acc(6′)-lb-cr*, *qnrA, qnrB, qnrC, qnrD, qnrS, qepA, oqxAB*) and *mcr* genes (*mcr-1* – *mcr-9*), respectively, by PCR as described previously (Zelendova et al., 2023; Janecko et al., 2018).

### 2.3. Whole-genome sequencing

We recovered at least one resistant *Escherichia* isolate from a total of 96 birds (67 out of 72 in 2018 and 28 out of 45 in 2019). Majority of *Escherichia* isolates (158 from 197 isolates) coming from the gulls and the water sample were subjected to short-read whole-genome sequencing (WGS) including all isolates from media with cefotaxime (99 isolates) with ESBL/AmpC production confirmed by MAST, all isolates with TMQR genes detected by PCR (33 isolates) and a representative subset of other ciprofloxacin-originated isolates without detected TMQR genes (25 from 62 isolates) and a single isolate with *mcr-1* from 3 *Escherichia* isolates obtained on media with colistin. The MAST and PCR testing was performed prior to the WGS to help us navigate the selection for WGS.

Genomic DNAs for WGS were isolated using NucleoSpin® Tissue kit (Macherey-Nagel GmbH & Co, Germany) and sequenced on Illumina (California, USA) platforms (MiSeq, HiSeq). A single isolate of the most prevalent sequence type belonging to a novel ST (11893) was selected for long-read sequencing using Sequel II platform (PacBio) as described previously (Nesporova et al., 2020).

### 2.4. WGS data processing and analysis

Raw sequencing data was quality ($Q \geq 20$) and adaptor trimmed using Trimmomatic software (v. 0.36) (Bolger et al., 2014). Assemblies were generated using SPAdes software (v. 3.11.1) (Bankevich et al., 2012). Fasta files were filtered for minimal a length of 200 bp. Sequence types (ST) were determined using MLST (v.2.0) (Larsen et al., 2012) and one isolate for each yet undescribed MLST allelic combination was uploaded to EnteroBase (Zhou et al., 2020; enterobase.warwick.ac.uk) to assign the novel ST (Supplementary Table 1). ResFinder (v.4.1) (Florensa et al., 2022), PlasmidFinder (v.2.1) (Carattoli et al., 2014) and VFDB (Chen et al., 2016) databases were employed via ABRicate

(github.com/tseemann/abricate) to evaluate the presence of antibiotic resistance genes (ARGs), plasmid replicons and virulence-associated genes (VAGs), respectively. Thresholds of 90 % for coverage and 90 % for identity were used for VAGs and plasmid replicons. Thresholds of 95 % for coverage and 95 % for identity were used to consider ARGs as present. Plasmid STs were assigned using pMLST (v. 2.0) (Carattoli et al., 2014). MGE (v. 1.0.3) (Johansson et al., 2021) was used to confer potential co-localozation of ARGs and plasmid replicons within the same contig. The phylogenetic group of *E. coli* isolates was determined by ClermonTyping (Beghain et al., 2018). The statistical analysis to compare the prevalence of cefotaxime-selected isolates in 2018 and 2019 and TMQR-positive isolates in 2018 and 2019 in gulls was done using Fisher's exact test ($\alpha = 0.05$). The relationship between individual ST persistence (defined as the ratio of isolates found in 2019 and 2018) and the number of VAGs was analysed and visualized using R package'utils' version 3.6.2. Alignment of contigs with $bla_{TEM}$ genes for *E. coli* ST2325 was performed using Geneious (v.7.1.9). The genomic data were deposited in GenBank within BioProjects PRJNA1025840 (*E. coli*) and PRJNA1028173 (*E. albertii*) and EnteroBase. The individual BioSamples and EnteroBase Uberstrain accession numbers are listed in Supplementary Table 1.

### 2.5. Phylogenetic analysis

All the phylogenetic analyses were based on Prokka open reading frames prediction (v1.14.5) (Seemann, 2014) and multi-fasta alignment using Roary (v.3.7.0) (Page et al., 2015) to infer the core genome (including genes present in 99 % or more genomes) for each specific set of strains. The trees were generated using RAxML (Stamatakis, 2014) supported by 100 bootstrap. After the first phylogenetic analysis of our 158 presumably *E. coli* isolates, we observed a distant clade comprising 17 of our isolates, represented mostly by previously undescribed ST12346. Moreover, the number of core genes included in the phylogenetic analysis was unusually low and comprised only of 2159 genes. We utilized the ClermonTyping results which suggested the strains could belong to *Escherichia albertii*. This was confirmed by additional phylogeny including all *E. albertii* in EnteroBase (564 strains, accessed 23rd June 2021) and our distant strains fitting among them (Supplementary Fig. 1). After these findings, we performed separate phylogenetic analysis for 141 *E. coli* strains (Fig. 1, Supplementary Table 6) and 17 *E. albertii* isolates (Fig. 2, Supplementary Table 10) which resulted in increasement to 2902 core genes for *E. coli*.

As four gulls from the Nove Mlyny reservoir were tagged with telemetry loggers, we could observe that Caspian gulls from this area migrate to several surrounding countries (Chytil et al., 2022). We wanted to investigate whether there are some overlaps between the isolates from this study and *E. coli* deposited in EnteroBase coming from these countries in terms of closely related strains. We downloaded a total of 5722 *E. coli* strains in June and July 2021 coming from Czechia (131 strains), Germany (2966, sampling year limited to 2010–2020), Poland (424 strains), Slovakia (35 strains), Austria (111 strains), Croatia (60), Hungary (195), the Netherlands (1024 strains, sampling year limited to 2010–2020) and Belgium (776). To further select the set of strains from EnteroBase and avoid the phylogeny being less specific because of including too distant strains, we ran the set through MLST (v. 2.0.9) (Larsen et al., 2012) and selected 1393 strains that had the same ST as at least one of our isolates. A phylogenetic tree for a total of 1534 strains (1393 from EnteroBase and 141 from this study) was generated and evaluated for SNP distance using snp-dists v.0.6.3 (github.com/tseemann/snp-dists) with results in Supplementary Table 9. We selected 75 strains from EnteroBase (Supplementary Table 2) which differed in 100 SNPs and less from our isolates and we rerun the phylogeny for these strains together with our *E. coli* strains (n = 141) to again evaluate the SNP distance for this phylogeny (Supplementary Table 7).

Separate phylogenetic analysis was performed for the strains of the most prevalent ST11893 (Supplementary Table 5) which was first detected in this study and for the second most prevalent ST2325 (Supplementary Table 8). The genome of *E. coli* RM48e of ST11893 was obtained as a hybrid assembly of PacBio polished data with Illumina reads using Unicycler (v.0.5.0) (Wick et al., 2017).

### 2.6. F34:A-:B- plasmid alignment

*E. coli* ST11893 was long-read-sequenced as a novel ST which showed to be the most prevalent among the whole set of gull isolates using a reference genome of *E. coli* RM48e. We have also obtained the full circular sequence of F34:A-:B- plasmid (GenBank accession number OR689867) with $bla_{CMY-2}$ using Unicycler with data from PacBio and Illumina. This reference plasmid sequence was used to align all ST11893 short-reads using BRIG (v.0.95) (Alikhan et al., 2011) and compare the plasmid sequences. BLASTn (blast.ncbi.nlm.nih.gov) was used to check whether there are some closely related plasmid sequences in GenBank with negative results.

### 2.7. Figures creation

The graphical abstract was created using Bio Render (biorender) and Google maps (google.com) images. Rstudio was used to visualize relation between persistence and average VAGs number for individual STs. The phylogenetic trees were visualised using the Interactive Tree of Life (iTOL) (Letunic and Bork, 2021). The F34:A-:B- alignment was visualised using BRIG (v.0.95). Annotated sequences were visualized in Geneious. MS PowerPoint was used to make schemes and adding labels within figures.

## 3. Results and discussion

### 3.1. Insights from dispersal behaviour and habitat selection of tracked Caspian gulls

The parallel study which tracked the movements of Caspian gulls from the colony in Nove Mlyny Reservoir showed that the gulls were mostly detected in lowlands areas of the Czech Republic, western Poland, Germany, the Netherlands and Belgium (Chytil et al., 2022). Closely related *E. coli* genomes were recovered from EnteroBase from all these countries. In several countries including Hungary, Slovakia, Austria and Croatia, gulls occurred less distinctively, yet closely related strains were still recovered from Hungary and Austria. Regarding Slovakia and Croatia, the low number of available genomes in EnteroBase could be the main reason for not observing a close phylogenetic link with our strains. The furthest dispersal distance of 1714 km of our gulls was recovered in Spain (Chytil et al., 2022), however, we did not involve Spain and other sporadically visited countries in our large phylogenetic analysis to keep it feasible.

Water bodies were the most frequently selected environment for our Caspian gulls followed by landfill sites, mineral extraction sites, and industrial and commercial units according to telemetry data (Chytil et al., 2022). The tracking of gulls showed that in the close area around the breeding colony, Caspian gulls frequently foraged on the water reservoirs, local landfill sites and nearby agricultural land. The nearest landfill site used by the gulls was situated 12 km from the breeding colony (Chytil et al., 2022). We assume it may be a likely source for the gulls' colony contamination with *E. coli* ST11893 (further details below) in 2018. Interestingly, Germany forbade to dump of untreated domestic waste in 2005 and after that, the gulls declined in population numbers with losing this important food source. This highlights the dual negative effect of anthropogenic influence as the landfill sites are not only already known hot-spots for AMR bacteria (enterobase.warwick.ac.uk) but also enable gulls to enlarge their population and inhabit more areas (Chytil et al., 2022; Klein and Neubauer, 2006).

Oppositely, several persistent STs, especially of *E. coli* ST2325 and *E. albertii* ST12346, seemed to be exchanging between gulls and water,

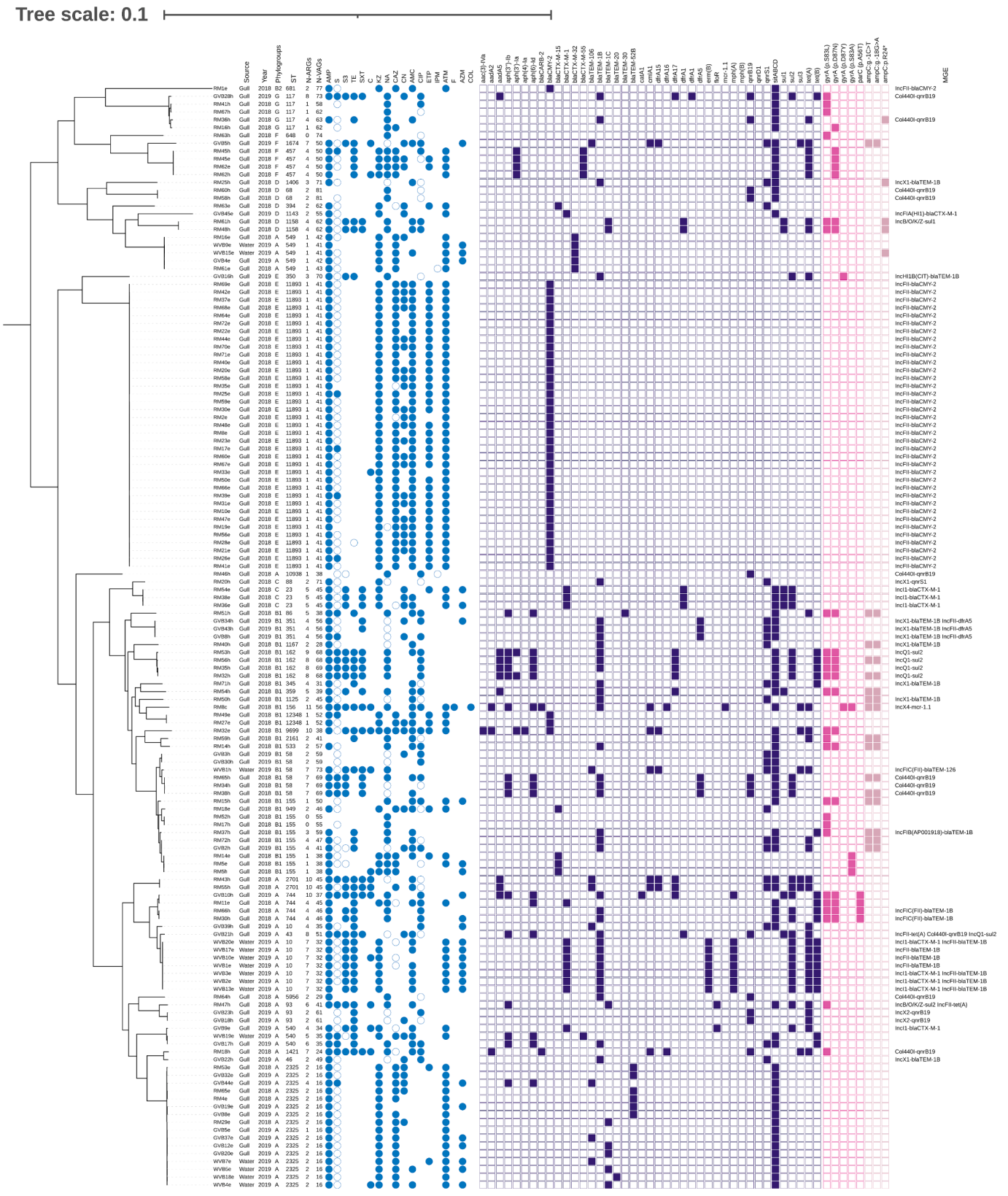


**Fig. 1.** The phylogenetic relationships of 141 *E. coli* isolates coming from this study (based on 2902 core genes). The columns in the first part represent the Source (gull or water), Year of isolation, Phylogroup, ST, total number of detected antibiotic resistance genes (N-ARGs), and total number of detected VAGs (N-VAGs). The columns with blue circles represent the results of phenotypic antibiotic resistance testing – full blue circle represent resistant phenotype, empty circle intermediate phenotype and no circle susceptible phenotype for respective antibiotics. The squares represent the presence (full) or absence (empty) of ARGs (violet), chromosomal mutations related to resistance (purple and pink). The last column shows which resistance genes were linked to be carried by specific plasmids (based on MGE tool). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

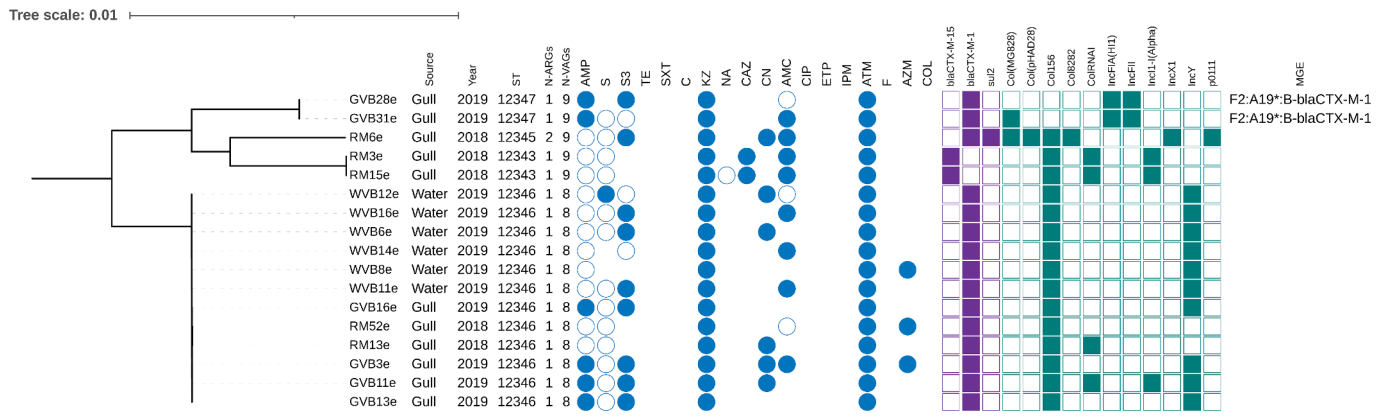


**Fig. 2.** The phylogenetic relationships of 17 *E. albertii* isolates coming from this study (based on 3671 core genes). The columns in the first part represent Source (gull or water), Year of isolation, ST, total number of detected antibiotic resistance genes (N-ARGs), and total number of detected VAGs (N-VAGs). The columns with blue circles represent the results of phenotypic antibiotic resistance testing – full blue circle represent resistant phenotype, empty circle intermediate phenotype and no circle susceptible phenotype for respective antibiotics. The squares represent the presence (full) or absence (empty) of ARGs (violet) and plasmid replicons (turquoise). The last column shows which resistance genes were linked to be carried by specific plasmids (based on MGE tool). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

as we recovered these STs from the water sample as well. However, with confluence of three rivers it is not possible to hypothesize about the primal source, considering also that the gulls themselves could be contaminated by these STs initially and introduced them to the water environment later.

This study showed valuable insights that could not be recovered from a snap-shot-based sampling of wild birds with no movement tracking available. There is still room for improvement for the design of these wildlife AMR surveillance studies including involving more sampling points, focusing on how long can be wild bird colonized by these AMR clones, what is the ability of these clones to spread in the bird's population, or identify the urban area at risk of pathogen spill over. Several studies focusing on these aspects have been already conducted showing that birds can carry ESBL-producing *E. coli* long enough to transfer them during migration and that interindividual transmission of these AMR clones among wild birds can be extensive (Sandegren et al., 2018), or showing that recreational urban water sources and schools are most sensitive to possible transmission of AMR bacteria from gulls in Madrid (Martín-Vélez et al., 2024).

### 3.2. *Escherichia* isolates obtained from different selective media in 2018 and 2019

We have obtained a total of 176 isolates from the cloacal samples of gulls and 21 isolates from the water sample. A total of 158 isolates comprising 141 *E. coli* and 17 *E. albertii* isolates were further evaluated by WGS. All the isolates were identified as *E. coli* by MALDI-TOF but later 17 cefotaxime-resistant (CefR) isolates were reclassified as *E. albertii* based on the analysis of WGS data. The limited MALDI-TOF ability to distinguish between *E. albertii* from *E. coli* may lead to its misclassification in previous studies and resulting in *E. albertii* being underreported. Therefore, next-generation sequencing or alternatively efficient method (e.g. additional biochemical tests) should be implemented to perform proper surveillance of this clinically relevant bacteria (Huys et al., 2003; Luo et al., 2021).

From the gull samples, we obtained CefR *Escherichia* isolates from 86 % (62/72) of samples in 2018 and 38 % (17/45) of samples in 2019 which was a significant decrease ($p < 0.00001$). In 2018, *E. coli* represented 57 of CefR isolates and 5 remaining were *E. albertii* and in 2019 there were 11 *E. coli* and 6 *E. albertii* CefR isolates. Ciprofloxacin-resistant (CipR) *E. coli* were collected from 81 % (58/72) and 80 % (36/45) of samples in 2018 and 2019, respectively. Three colistin-resistant (ColR) *E. coli* isolates were recovered in 2018 only. Regarding the *Escherichia* isolates coming from the water sample, we obtained the intended number of 20 *E. coli*-like colonies from media with cefotaxime but only one from media with ciprofloxacin. No isolate was obtained from colistin-supplemented media.

PCR screening for TMQR genes in 94 CipR isolates from gulls revealed a slight and statistically insignificant ($p < 0.2$) increase in the prevalence of these genes among gulls from 2018 with 33 % (16/58) to 2019 with 41 % (15/36). In 2018, 21 % of CipR isolates from gulls carried *qnrB* and 12 % *qnrS* and opposite to that in 2019 *qnrS* was present in 30 % of CipR isolates and *qnrB* in 11 %. Gene *qnrS* was also detected in a single CipR isolate from the water sample. The WGS revealed all *qnrB* being *qnrB19* variant and *qnrS* as *qnrS1*. From three ColR *E. coli* only one carried *mcr-1* gene. In *E. coli* set (141), we observed a high prevalence of $bla_{\text{CMY-2}}$ (29 %, n = 141), $bla_{\text{TEM-1B}}$ (26 %), $bla_{\text{CTX-M-1}}$ (9 %), $bla_{\text{TEM-106}}$ and $bla_{\text{TEM-1C}}$ (both 6 %). Interestingly, we detected $bla_{\text{CTX-M}}$ genes in all *E. albertii* isolates represented by $bla_{\text{CTX-M-1}}$ (15 isolates) and $bla_{\text{C-15}}$ (2 isolates), as the only ARGs present in all except one isolate (Fig. 2). The results of phenotypic testing of resistance to 18 antibiotics are summarized in Fig. 1 for *E. coli* and Fig. 2 for *E. albertii*. The phenotypic resistance is summarized and compared to WGS data in the Supplementary results.

### 3.3. Phylogenetic analysis and persistence evaluation

One of the most important benefits of this study is that we can compare the diverse *Escherichia* strains presence in the same gulls' colony within twelve months' period. We also provided information on which *Escherichia* isolates were present in the water environment in 2019. Moreover, we compared our data with EnteroBase so we can provide a wider perspective and determine potential regional high-risk STs related to these wandering birds. Therefore, we can provide insights into the dissemination and persistence of specific *Escherichia* lineages.

The most abundant phylogenetic groups among isolates from this study were A (n = 46 isolates), E (n = 38) and B1 (n = 34) which are related to the typically non-pathogenic *E. coli* (Cummins et al., 2022). However, we have also detected several isolates from more likely pathogenic phylogroups D (n = 7 isolates), F (n = 6), G (n = 5) and B2 (n = 1).

A total of 41 different *E. coli* STs were detected among our isolates and several of them were found in both years or overlapping between the gulls and a water sample. In the phylogenetic analysis of all of our *E. coli* isolates (involving 2902 core genes), four unique water-gull

isolate pairs showed differences up to 10 SNPs including ST2325 and ST540 (Supplementary Table 6). However, in the detailed phylogeny of ST2325 (involving 3394 core genes), the closest pair of gull and water isolates differed by 17 SNPs (Supplementary Table 8). Closely related strains (up to 100 SNPs) of detected STs were also recovered from EnteroBase, specifically ST744 (n = 26), ST23 (n = 17), ST949 (n = 11), ST155, ST58 (both n = 4), ST162 (n = 3), ST117, ST9699, ST86 (each n = 2), ST10, ST350, ST648, ST1167 (each n = 1), highlighting their ability to circulate across various sources in Czechia and surrounding countries (Fig. 3, Supplementary Fig. 2). We further focus on notable STs regarding their prevalence, persistence and dissemination traits.

### 3.4. Temporary dominant local clone of E. coli ST11893

The ST11893 was a novel ST first detected in this study and all ST11893 isolates carried *bla*$_{CMY-2}$ gene incorporated within IncFII F34: A-:B- plasmid. ST11893 was remarkably prevalent among *Escherichia* isolates selected from 2018 using cefotaxime media as it represented 59.7 % (37/62) of recovered CefR isolates. Moreover, as the level of samples from which were recovered cefotaxime-resistant *Escherichia* isolates surprisingly dropped from 86 % in 2018 to 38 % in 2019, we can assume that gulls were temporarily colonized with this particular ST. The isolates were mostly clonal in the pairwise comparison with the medium of 8 SNPs distance and a maximum of 31 SNPs (Supplementary Table 5) which suggest they came from a common source. The isolates

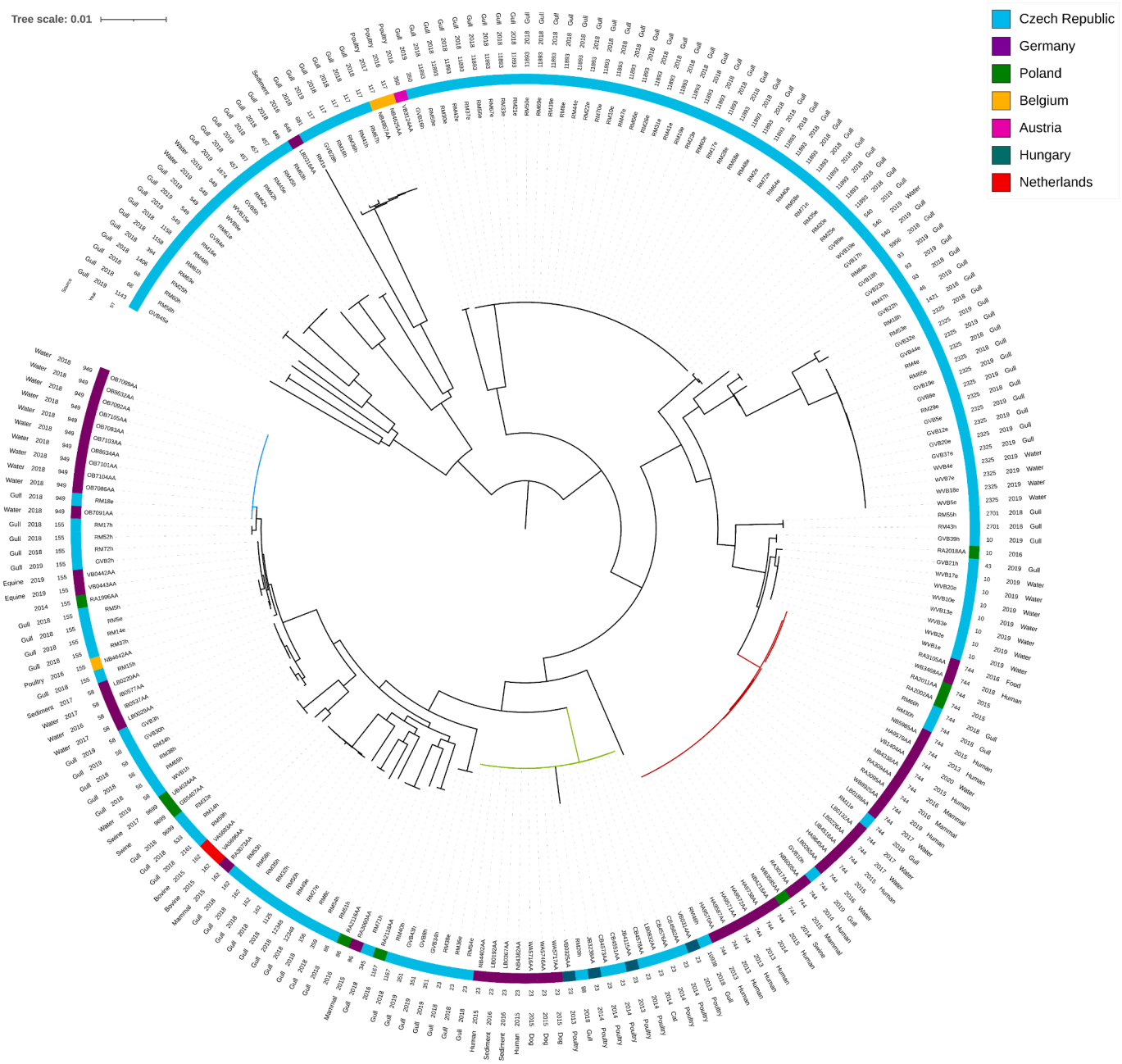


**Fig. 3.** The phylogenetic relationships of *E. coli* isolates from this study and 75 *E. coli* isolates coming from EnteroBase that were related up to 100 SNPs with at least one of our samples in the preliminary phylogeny with 54 remained up to 100 SNPs within this detailed phylogeny (2849 core genes). The tree branches with the highest abundance of isolates from EnteroBase are highlighted in colour: ST744 (red), ST23 (green), ST949 (blue). The colour in the inner circle represents the country of origin (see the figure legend), the next circles represent ST, the year of isolation and source of isolation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

reached an average number of 42 VAGs which is not extensive VAGs cargo for *E. coli* as some strains can carry over 100 VAGs (Jarocki et al., 2020). As we did not detect a single isolate of this ST the next year, despite the same sampling spot, the potential explanation could be that the ST11893 was overruled by other strains as a result of bacterial competition or the dismissal of the temporary source of contamination. We currently study the influence of bacterial competition in this case in a follow-up study. Moreover, not a single isolate from 2019 carried the IncFII F34:A-:B- plasmid associated with ST11893.

### 3.5. Locally persistent lineages of *E. coli* ST2325 and *E. albertii* ST12346

With 16 isolates overall, *E. coli* ST2325 was the second most prevalent (11 %; 16/141) from the sequenced *E. coli* collection. In contrast to ST11893, it was detected in gulls from both years (2018 and 2019), and also from the water sample so it seems to be more persistent and omnipresent in this specific location. Similarly, with 12 isolates, *E. albertii* ST12346 was the third most prevalent ST detected for the whole set and the most prevalent for *E. albertii*.

Interestingly, a common trait for ST2325 and ST12346 was a low number of VAGs, as ST2325 has the lowest number of VAGs from *E. coli* with 16 VAGs and ST12346 has the lowest number in total with only 8 VAGs on average. Our observation of presumably low-virulent *E. albertii* is in contrast with its recent genomic study which emphasizes its high virulence content and pathogenic potential (Luo et al., 2021).

The relationship between ST persistence (defined as the ratio of isolates found in 2019 and 2018 for respective ST) and the average number of VAGs in the given ST is shown in Supplementary Figure 5. The qualitative overview of present VAGs in *E. coli* isolates is shown in Supplementary Figure 6 and in *E. albertii* in Supplementary Figure 7. Our results suggest that higher number of VAGs might not be helpful for environmental persistence of an *Escherichia* ST, rather the opposite. However, this observation is limited because of a low number of total isolates and by the study design as the water sample was taken only in 2019. Therefore, the analysis reflects on the combination of both, the ability of ST to persist within the gulls per se and the ability of ST to survive in the water. As clonality was not generally observed for ST2325 (presented in the Supplementary results), even for some isolates from the same year, it is probable that this ST occurred for the first time in the location longer time prior our sampling as it has the opportunity to explore multiple evolution trajectories. The traits enabling *Escherichia* to survive in water for prolonged time should be studied further as the circulation of environmentally successful ARB between water and wild birds is unsettling. Special focus should be paid on the role of VAGs in this as another study identified that *E. coli* from a contaminated surface water have low virulent content as well and it has been shown already reported that bacterial isolates can show reduced virulence as a result of avoiding the phage predation and developing resistance to them (León and Bastías, 2015).

The ability to persist is most certainly multifactorial but our data put the VAGs into novel perspective as they are generally seen as something which should make *E. coli* more prone to succeed (Manges et al., 2019). However, that could be related to the common bias toward clinical isolates as they are studied heavily but still represent only a specific subgroup of *E. coli* (Wyrsch et al., 2022; Torres et al., 2020; Manges et al., 2019 Jun 12).

### 3.6. Commonly detected sequence types

Here, we would like to highlight *E. coli* STs that were detected in both years (ST117, ST744, ST93, ST58, ST155, ST549) or they came from both, gulls and the water sample (ST10, ST540, ST58, ST549).

These STs were also common among the 75 strains recovered from EnteroBase which were closely related (up to 100 SNPs in the phylogeny analysis of 1534 strains, with 2138 core genes, Supplementary Table 9) to ours from the geographically relevant area (Chytil et al., 2022). We detected 13 different STs among the 75 strains, 26 strains belonged to ST744 and 17 to ST23. ST744 strains came from 2013 to 2020 with origin mainly in Germany and also in Poland. The minimum SNPs difference for ST744 for a pair of isolates from our study (gull) and EnteroBase (human, Germany) was 35 (preliminary phylogeny, 2138 core genes) or 46 (final phylogeny, 2849 core genes, described in the Supplementary results). ST23 came from 2013 to 2018, it was found in poultry and a cat from Czechia and in various sources in Hungary and Germany as well with a minimum of 52 SNPs difference (sediment, Germany) from our isolate (in both preliminary and final phylogeny). Other represented STs in the EnteroBase subset included the common ones such as ST10, ST58, ST117, ST155, ST162, ST648 and also rarer STs including ST350, ST949, ST1167 and ST9699. Nevertheless, these rare STs could be underreported because they could only be evolutionary younger and it does not necessarily mean they are harmless. For example, ST9699 from a gull carried the highest number of ARGs from the gull collection with eleven genes detected and this ST can be found in wildlife and livestock in Poland. We detected only one isolate (ST949) in our set which was clonally related to genomes from EnteroBase linked to German surface waters (further details in the Supplementary results).

### 3.7. F34 plasmids

F34:A-:B- plasmids were the most prevalent in the collection due to their association with the most prevalent *E. coli* ST11893. We aligned all these plasmids with the reference plasmid obtained through the combination of long- (PacBio) and short-reads and we saw all possessed the identical structure (Fig. 4). Interestingly, while looking closer at the surrounding of $bla_{CMY-2}$, we observed that the insertion of this resistance gene led to the truncation of colicin-related *cba* gene (Fig. 4). Colicins are polypeptide toxins produced by and active against *E. coli* and closely related bacteria so they are an important tool for the inner competition among these bacteria (Cascales et al., 2007). We hypothesize that this truncation could be one of the factors which contributed to the disappearance of the ST11893 and we intend to conduct further investigation on this lineage and its F34:A-B- plasmid in further studies. It would be an interesting example that acquiring a resistance gene may turn into a disadvantage in long-term lineage environmental persistence and survival. The other notable plasmids and their association with ARGs are presented in the Supplementary results.

## 4. Conclusions

We showed that *E. coli* isolates closely related to the ones obtained within our study can be detected in countries to which these particular gulls wander with a particularly high proportion of closely related ST744 and ST23. We detect an unexpected decrease of extended-spectrum beta-lactam-resistant *E. coli* from 2018 (86 %) to 2019 (38 %) among gulls. However, we were able to explain this by the high occurrence of ST11893-$bla_{CMY-2}$ clone in 2018 which did not persist among the gulls to the following year. This is an example that antibiotic resistance itself may not always be the key for a lineage to survive, especially outside the clinical environment. On the other hand, we also noted that some STs managed to persist in the sampled gulls' population into the following year and some were also detected in the water close to the gulls' colony. Here, we observed that the lineages of ST2325 (*E. coli*) and ST12346 (*E. albertii*) with the best-observed persistence had also the lowest number of VAGs from *E. coli* and *E. albertii*.

## Funding

The study was supported by the Internal Grant Agency of University of Veterinary Sciences Brno (Grant Number 209/2023/FVHE) and the Ministry of Health of the Czech Republic (Grant Number NU20J-09-00040).

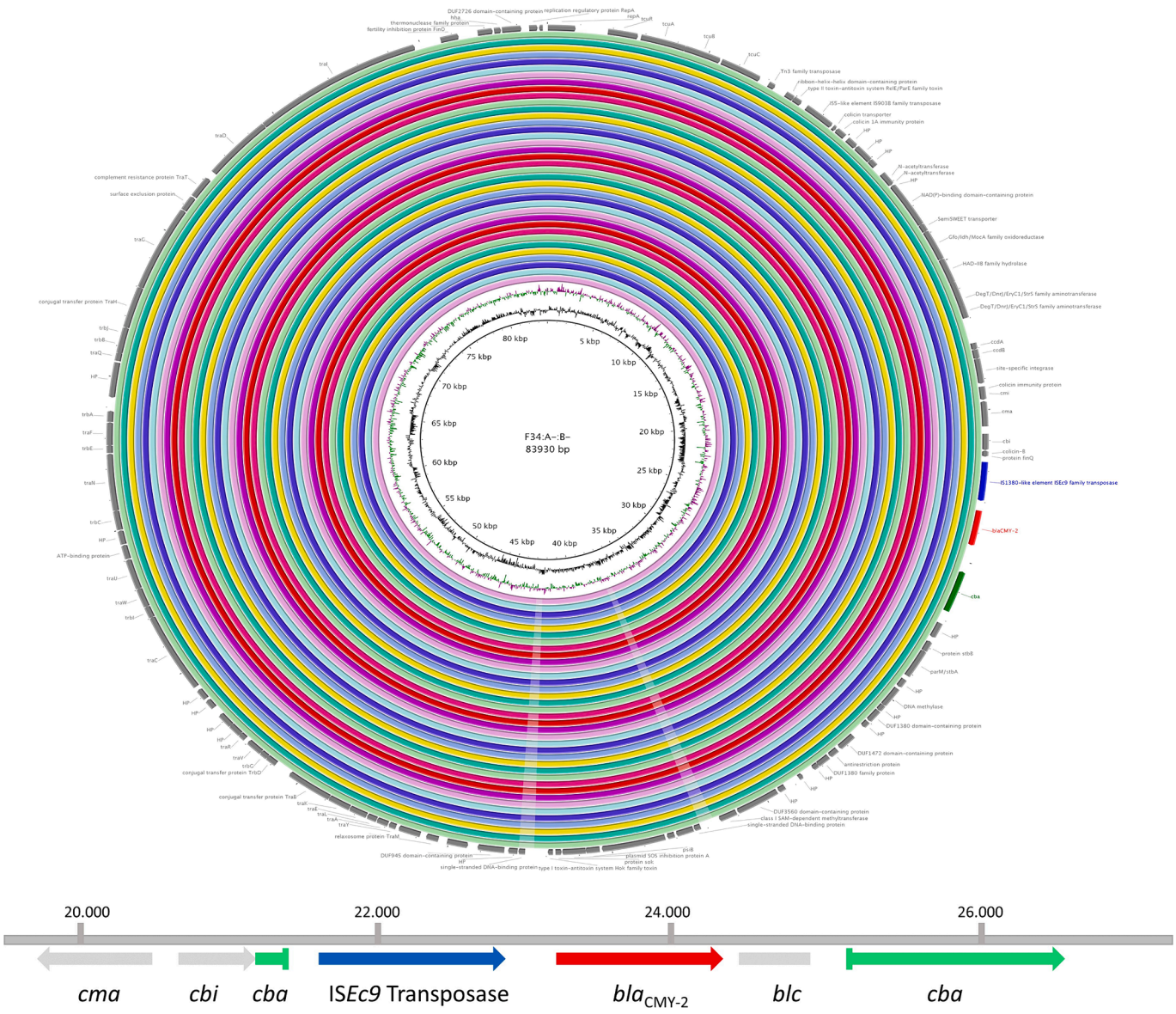


**Fig. 4.** The alignment of F34:A-:B- plasmids from ST11983 isolates. The environment of $bla_{CMY-2}$ (red) gene is depicted in detail showing the insertion of this resistance gene via IS*Ec9* (blue) caused truncation of colicin production-related gene *cba* (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

## CRediT authorship contribution statement

**Kristina Nesporova:** Writing – original draft, Visualization, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Michaela Ruzickova:** Writing – review & editing, Investigation, Data curation. **Hassan Tarabai:** Writing – review & editing, Project administration, Data curation. **Simon Krejci:** Writing – review & editing, Data curation. **Martina Masarikova:** Investigation. **Jarmila Lausova:** Writing – review & editing, Investigation, Data curation. **Ivan Literak:** Conceptualization. **Monika Dolejska:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgements

Part of this work was presented at the following conferences: 32th European Congress of Clinical Microbiology and Infectious Diseases (ECCMID), 2022, Lisbon, Portugal and One Health EJP Scientific Meeting, 2021, Copenhagen, Denmark.

We thank V. Palsovicsova, P. Sismova, M. Klvana, K. Rehakova, I. Kutilova, A. Valcek, D. Cervinkova, Z. Schovancova, J. Palkovicova, T. Nohejl, and I. Jamborova for their help in the laboratory and in the field. We thank I. Bitar and J. Hrabak for their help with long-read sequencing. We acknowledge institutions that share their WGS data in EnteroBase and other publicly available repositories.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2024.108606.

## References


Ahlstrom, C.A., van Toor, M.L., Woksepp, H., et al., 2021. Evidence for continental-scale dispersal of antimicrobial resistant bacteria by landfill-foraging gulls. Sci. Total Environ. 764, 144551 https://doi.org/10.1016/j.scitotenv.2020.144551.

Ahmed, Z.S., Elshafiee, E.A., Khalefa, H.S., Kadry, M., Hamza, D.A., 2019. Evidence of colistin resistance genes (mcr-1 and mcr-2) in wild birds and its public health implication in Egypt. Antimicrob. Resist. Infect. Control 8 (1), 1–8. https://doi.org/10.1186/s13756-019-0657-5.

Alikhan, N.F., Petty, N.K., Ben Zakour, N.L., Beatson, S.A., 2011. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. BMC Genomics 12. https://doi.org/10.1186/1471-2164-12-402.

Atterby, C., Ramey, A.M., Hall, G.G., Järhult, J., Börjesson, S., Bonnedahl, J., 2016. Increased prevalence of antibiotic-resistant *E. coli* in gulls sampled in Southcentral Alaska is associated with urban environments. Infect Ecol Epidemiol. 6 (1) https://doi.org/10.3402/IEE.V6.32334.

Atterby, C., Börjesson, S., Ny, S., Järhult, J.D., Byfors, S., Bonnedahl, J., 2017. ESBL-producing *Escherichia coli* in Swedish gulls—a case of environmental pollution from humans? PLoS One 12 (12), 1–13. https://doi.org/10.1371/journal.pone.0190380.

Bankevich, A., Nurk, S., Antipov, D., et al., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19 (5), 455–477. https://doi.org/10.1089/cmb.2012.0021.

Beghain, J., Bridier-Nahmias, A., Le, N.H., Denamur, E., Clermont, O., 2018. ClermonTyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping. Microb. Genomics 4 (7), 1–8. https://doi.org/10.1099/mgen.0.000192.

biorender.com.

https://blast.ncbi.nlm.nih.gov/Blast.cgi.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30 (15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Carattoli, A., Zankari, E., Garciá-Fernández, A., et al., 2014. In silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. Antimicrob. Agents Chemother. 58 (7), 3895–3903. https://doi.org/10.1128/AAC.02412-14.

Cascales, E., Buchanan, S.K., Duché, D., et al., 2007. Colicin biology. Microbiol. Mol. Biol. Rev. 71 (1), 158–229. https://doi.org/10.1128/mmbr.00036-06.

Chen, L., Zheng, D., Liu, B., Yang, J., Jin, Q., 2016. VFDB 2016: Hierarchical and refined dataset for big data analysis - 10 years on. Nucleic Acids Res. 44 (D1), D694–D697. https://doi.org/10.1093/nar/gkv1239.

Chytil, J., Krejčí, Š., Rozsypalová, L., Literák, I., 2022. History, dispersal and habitat selection in Central European Caspian gulls *Larus cachinnans*: a study of birds originating from a single breeding colony. Biologia (bratisl). 77 (2), 395–409. https://doi.org/10.1007/s11756-021-00938-z.

Clinical and Laboratory Standards Institute (CLSI). Performance Standards for Antimicrobial Susceptibility Testing. 27th ed. CLSI supplement M100 (ISBN 1-56238-804-5 [Print]; ISBN 1-56238-805-3 [Electronic]). Clinical and Laboratory Standards Institute, 950 West Valley Road, Suite 2500, Wayne, Pennsylvania 19087 USA, 2017.CLSI: Clinical and Laboratory Standards Institute. Performance standards for antimicrobial susceptibility testing; twenty-first informational supplement. CLSI document M100-S21. CLSI, Wayne, PA, USA, 2011.

The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 10, 2020. http://www.eucast.org.

Cummins, E.A., Hall, R.J., Connor, C., McInerney, J.O., McNally, A., 2022. Distinct evolutionary trajectories in the *Escherichia coli* pangenome occur within sequence types. Microb. Genomics 8 (11), 1–13. https://doi.org/10.1099/mgen.0.000903.

https://enterobase.warwick.ac.uk/.

Florensa, A.F., Kaas, R.S., Clausen, P.T.L.C., Aytan-Aktug, D., Aarestrup, F.M., 2022. ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. Microb. Genomics 8 (1), 1–10. https://doi.org/10.1099/mgen.0.000748.

https://github.com/tseemann/abricate.

https://github.com/tseemann/snp-dists.

https://www.google.com/maps.

Huys, G., Cnockaert, M., Janda, J.M., Swings, J., 2003. *Escherichia albertii* sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. Int. J. Syst. Evol. Microbiol. 53 (3), 807–810. https://doi.org/10.1099/ijs.0.02475-0.

Janecko, N., Halova, D., Jamborova, I., et al., 2018. Occurrence of plasmid-mediated quinolone resistance genes in *Escherichia coli* and *klebsiella* spp. recovered from *Corvus brachyrhynchos* and *Corvus corax* roosting in Canada. Lett. Appl. Microbiol. 67 (2), 130–135. https://doi.org/10.1111/lam.12993.

Jarocki, V.M., Reid, C.J., Chapman, T.A., Djordjevic, S.P., 2020. *Escherichia coli* ST302: genomic analysis of virulence potential and antimicrobial resistance mediated by Mobile genetic elements. Front. Microbiol. 10 (January), 1–17. https://doi.org/10.3389/fmicb.2019.03098.

Johansson, M.H.K., Bortolaia, V., Tansirichaiya, S., Aarestrup, F.M., Roberts, A.P., Petersen, T.N., 2021. Detection of mobile genetic elements associated with antibiotic resistance in *Salmonella enterica* using a newly developed web tool: MobileElementFinder. J. Antimicrob. Chemother. 76 (1), 101–109. https://doi.org/10.1093/JAC/DKAA390.

Klein, R., Neubauer, G., 2006. Influxes of Caspian gulls Larus cachinnans and yellow-legged Gulls L. michahellis into northern Central Europe – origin, causes, course and trend. Vogelwelt 127, 91–97.

Larsen, M.V., Cosentino, S., Rasmussen, S., et al., 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J. Clin. Microbiol. 50 (4), 1355–1361. https://doi.org/10.1128/JCM.06094-11.

León, M., Bastías, R., 2015. Virulence reduction in bacteriophage resistant bacteria. Front. Microbiol. https://doi.org/10.3389/fmicb.2015.00343.

Letunic, I., Bork, P., 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49 (W1), W293–W296. https://doi.org/10.1093/nar/gkab301.

Liu, L., Wang, S., Chen, J., 2021. Transformations from specialists to generalists cause bacterial communities are more stable than micro-eukaryotic communities under anthropogenic activity disturbance. Sci. Total Environ. 790, 148141 https://doi.org/10.1016/j.scitotenv.2021.148141.

Luo, L., Wang, H., Payne, M.J., et al., 2021. Comparative genomics of chinese and international isolates of *Escherichia albertii*: population structure and evolution of virulence and antimicrobial resistance. Microb Genomics 7 (12). https://doi.org/10.1099/MGEN.0.000710.

Manges, A.R., Geum, H.M., Guo, A., Edens, T.J., Fibke, C.D., Pitout, J.D.D., 2019 Jun 12. Global extraintestinal pathogenic Escherichia coli (ExPEC) lineages. Clin. Microbiol. Rev. 32 (3), e00135–e00218. https://doi.org/10.1128/CMR.00135-18.

Martín-Vélez, V., Navarro, J., Figuerola, J., Aymí, R., Sabaté, S., Planell, R., Vila, J., Montalvo, T., 2024. A spatial analysis of urban gulls contribution to the potential spread of zoonotic and antibiotic-resistant bacteria. Sci. Total Environ. 20 (912), 168762 https://doi.org/10.1016/j.scitotenv.2023.168762.

Na, G., Lu, Z., Gao, H., et al., 2018. The effect of environmental factors and migration dynamics on the prevalence of antibiotic-resistant *Escherichia coli* in estuary environments. Sci. Rep. 8 (1), 1–9. https://doi.org/10.1038/s41598-018-20077-x.

Nesporova, K., Wyrsch, E.R., Valcek, A., Bitar, I., Chaw, K., Harris, P., Hrabak, J., Literak, I., Djordjevic, S.P., Dolejska, M., 2020. *Escherichia coli* sequence type 457 is an Emerging extended-spectrum-β-lactam-resistant lineage with reservoirs in wildlife and food-producing animals. Antimicrob. Agents Chemother. 65 (1), e01118–e01120. https://doi.org/10.1128/AAC.01118-20.

Page, A.J., Cummins, C.A., Hunt, M., et al., 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31 (22), 3691–3693. https://doi.org/10.1093/bioinformatics/btv421.

Sandegren, L., Stedt, J., Lustig, U., Bonnedahl, J., Andersson, D.I., Järhult, J.D., 2018. Long-term carriage and rapid transmission of extended spectrum beta-lactamase-producing E. coli within a flock of Mallards in the absence of antibiotic selection. Environ. Microbiol. Rep. 10 (5), 576–582. https://doi.org/10.1111/1758-2229.12681.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30 (14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

Sismova, P., Sukkar, I., Kolidentsev, N., Palkovicova, J., Chytilova, I., Bardon, J., Dolejska, M., Nesporova, K., 2023. Plasmid-mediated colistin resistance from fresh meat and slaughtered animals in the Czech Republic: nation-wide surveillance 2020–2021. Microbiol. Spectr. 11 (5), e0060923. https://doi.org/10.1128/spectrum.00609-23.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30 (9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

R. Torres, J. Carvalho, M. Cunha, E. Serrano, J. Palmeira, C. Fonseca, A systematic review of AMR in wildlife: temporal and geographical distribution trends. 2020:1-15. https://doi.org/10.22541/au.159466958.84737751.

Van Puyvelde, S., Deborggraeve, S., Jacobs, J., 2018. Why the antibiotic resistance crisis requires a one health approach. Lancet Infect. Dis. 18 (2), 132–134. https://doi.org/10.1016/S1473-3099(17)30704-1.

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E., 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput. Biol. 13 (6), 1–22. https://doi.org/10.1371/journal.pcbi.1005595.

Wyrsch, E.R., Nesporova, K., Tarabai, H., et al., 2022. Urban wildlife crisis: Australian silver gull is a bystander host to widespread clinical antibiotic resistance. mSystems. 7 (3), 1–16. https://doi.org/10.1128/msystems.00158-22.

Zelendova, M., Papagiannitsis, C.C., Sismova, P., et al., 2023. Plasmid-mediated colistin resistance among human clinical *enterobacterales* isolates: national surveillance in the Czech Republic. Front. Microbiol. 14 (April), 1–13. https://doi.org/10.3389/fmicb.2023.1147846.

Zhou, Z., Alikhan, N.F., Mohamed, K., Fan, Y., Achtman, M., 2020. The EnteroBase user's guide, with case studies on *Salmonella transmissions*, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. Genome Res. 30 (1), 138–152. https://doi.org/10.1101/gr.251678.119.