One Health 20 (2025) 100970

Contents lists available at ScienceDirect

One Health

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

Subtle genomic differences in *Klebsiella pneumoniae sensu stricto* isolates indicate host adaptation

Bridget O'Brien [a], Alla Yushchenko [a], Jinha Suh [a], Dongyun Jung [a,b,c], Zhangbin Cai [a], Ngoc Sang Nguyen [a], Makeda Semret [d], Simon Dufour [b,e,c], Séamus Fanning [f,g], Jennifer Ronholm [a,b,c,*]

[a] Faculty of Agricultural and Environmental Sciences, Macdonald Campus, McGill University, Montreal, Quebec, Canada
[b] Mastitis Network, Saint-Hyacinthe, Quebec, Canada
[c] Regroupement FRQNT Op+Lait, Saint-Hyacinthe, Quebec, Canada
[d] McGill University Health Centre (Infectious Diseases and Medical Microbiology), Montreal, Quebec, Canada
[e] Faculté de Médecine Vétérinaire, Université de Montréal, Saint-Hyacinthe, Quebec, Canada
[f] UCD-Centre for Food Safety, Science Centre South, University College Dublin, Dublin, Ireland
[g] Institute for Global Food Security, School of Biological Sciences, Queen's University Belfast, 19 Chlorine Gardens, Belfast BT9 5DL, Ireland

ARTICLE INFO

Keywords:
Klebsiella pneumoniae sensu stricto
Comparative genomics
Opportunistic pathogen
Mastitis

ABSTRACT

Klebsiella pneumoniae sensu stricto (KpI) is an opportunistic pathogen capable of residing as a commensal in both human and bovine intestinal tracts and can cause serious systemic infections in humans and severe clinical mastitis in dairy cattle. It is unclear what role zoonotic and anthroponotic transmission play in the dissemination of KpI. In this study, we use a comparative genomic approach to identify differences between KpI associated with disease in humans and cattle and aimed to identify any potential genetic barriers limiting transmission of KpI between these two hosts. A total of 128 KpI strains (bovine $n = 65$; human $n = 63$) were whole genome sequenced and human and bovine strains were compared based on phylogenomics, the pangenome, mobile genetic elements, and differential gene abundance. No obvious phylogenomic differentiation was observed between isolates from these hosts. However, subtle genetic differences exist between bovine and human KpI which likely reflect environmental adaptation to different host niches, including a higher representation of gene clusters encoding ferric citrate uptake transporters, as well as histidine, arginine, and lactose utilization pathways in bovine isolates. These gene clusters may be positively selected due to the unique metabolic environment of the mammary gland, where lactose, citrate-bound iron, and amino acids like histidine and arginine provide growth advantages for KpI during mastitis. Overall, our study identified no obvious genetic barriers to zoonotic transmission of KpI within the dairy environment and provides insight into the development of host-specific therapeutic options for KpI infections in humans and bovine.

1. Introduction

Klebsiella pneumoniae species complex (KpSC) is ubiquitous in nature, commonly found in both animals and humans, and can pose a serious threat to public health due to its propensity towards antimicrobial resistance (AMR) [1]. In humans, the carriage rate of *K. pneumoniae sensu stricto* (KpI) in the gastrointestinal tract ranges from 6 % to 55.9 %, with variance attributed to age, location, travel history, or recent antibiotic use [2–4]. Although commonly found as a commensal, KpSC can act as an opportunistic pathogen that primarily affects immunocompromised individuals and may result in severe infections [5]. In humans, KpI is a leading cause of hospital-acquired infections, including urinary tract infections, pneumonia, and bacteremia [6]. In cattle, particularly on dairy farms, *Klebsiella spp.* are isolated from various sources such as rumen samples, water, bedding, alleyways, holding pens, and soil [7]. Among dairy cattle, *Klebsiella spp.* is recognized to cause severe clinical mastitis, resulting in economic losses for the dairy industry due to reduced milk production, culling, and/or death [6,8].

The expression of specific virulence factors by KpSC can influence

* Corresponding author at: Faculty of Agricultural and Environmental Sciences, Macdonald Campus, McGill University, Montreal, Quebec, Canada.
E-mail address: Jennifer.Ronholm@mcgill.ca (J. Ronholm).

https://doi.org/10.1016/j.onehlt.2025.100970
Received 19 June 2024; Received in revised form 15 November 2024; Accepted 7 January 2025
Available online 11 January 2025

infection severity. The capsule is an integral structural component that protects KpSC from host immune responses [9]. Siderophore systems encoded by KpSC, such as aerobactin, salmochelin, yersniabactin, and enterobactin each contribute to iron acquisition that is essential for growth and replication, especially in environments where iron concentrations are low, such as the bovine mammary gland [9,10]. The influence of each siderophore system on virulence varies, with aerobactin associated with high virulence and enterobactin associated with lower virulence in humans, likely due to their different affinities for iron [9]. Hypervirulent strains of KpI, associated with capsule serotype K1 or K2, aerobactin and/or yersiniabactin siderophore systems, and regulators of a hypermucoid phenotype, are particularly concerning as they can cause invasive infection in otherwise healthy individuals [11].

KpSC is also known for its high abundance and repertoire of AMR genes, owing to its unusual ability to acquire and maintain a diverse array of plasmids [12]. Use of antibiotics in agriculture can select for AMR pathogens which may transfer to the environment through various means [12]. While the World Health Organization does not classify KpI as a zoonotic pathogen, some argue that its zoonotic potential might be underestimated [13]. Agricultural workers directly exposed or KpI spread may occur indirectly through consumption of contaminated food or water [14]. However, little is known about the genomic relationship between KpSC of bovine and human origins, and whether bovine KpSC represents a significant threat to human health. Recent studies have demonstrated that the *lac* and *fec* operons are significantly associated with bovine KpI, but little is known otherwise as most studies focus on human strains rather than those of animal origin [1,15]. Therefore, further research is needed to understand host-adaption of KpSC and identify genomic barriers that would limit transmission between hosts.

In this study, we aimed to perform comparative genomics on 128 KpI isolates from both human and bovine infections. We characterized the pangenome of the KpI isolates, provided insight into the evolutionary relationship between bovine and human strains, and compared their genomic content – providing detailed insight into host-adaptation.

2. Materials and methods

2.1. Bacterial isolates, sequencing, and assembly

A total of 128 KpI isolates were used in this study. Bovine isolates, collected between 2007 and 2008, were part of the Mastitis Pathogen Culture Collection [16]. Human isolates were cultured from blood samples taken from patients at McGill University Health Centre in 2020, following institutional ethical guidelines set by McGill University and the Declaration of Helsinki, as described previously [17]. Whole genome sequencing and assembly of KpI isolates was described previously [17].

2.2. Gene annotation & pangenome analysis

Species, sequence type (ST), K-locus type, and O-locus type determination as well as the identification of common iron acquisition systems, were completed using Kleborate (v.2.4.0) [18]. The K- and O-locus types are accompanied with a confidence level of either perfect, very high, high, good, low, or none, as described here [19]. All genome assemblies served as input into Prokka (v.1.14.5), generating gff files which were used as input into Panaroo (v.1.3.3) [20,21]. Panaroo was run using 'sensitive' mode and MAFFT to align the core genome of KpI [22]. Gene accumulation curves were generated using the specaccum function from the R package, vegan (v.2.6–4), on a binary version of the gene presence/absence file generated by Panaroo [23]. The core gene alignment of KpI was used as input into FastTree (v.2.1.11) using default parameters, generating an approximately-maximum-likelihood phylogenetic tree [24]. The tree was visualized and annotated in R using ggtree package (v.3.10.0) [25].

2.3. Identification of mobile genetic elements

Plasmid replicon markers were identified using Abricate and the PlasmidFinder Database with default parameters (v.1.0.0) [26,27]. Mlplasmids (v.2.1.0), a machine-learning classifier designed to predict whether short-read contigs in *K. pneumoniae* are plasmid- or chromosome- derived, was used to identify if contigs containing replicons identified by PlasmidFinder were predicted to originate from plasmids rather than the chromosome [28]. Contigs predicted to be from plasmids were used in downstream analyses. Statistical significance was assessed using a Fisher's exact test in Scoary (v.1.6.16) which utilized a plasmid presence/absence file as input [29]. From this analysis, the Benjamini Hochberg adjusted *p*-value and odds ratio (OR) were recorded.

2.4. Differential gene analysis

The presence of mucoid phenotype regulators (*rmp* locus, *rmpA2*) and iron acquisition systems (aerobactin, enterobactin, salmochelin, and yersiniabactin) with respective lineage, was determined using Kleborate [18]. Genes encoding ferric citrate iron uptake systems were identified in Panaroo output. Significance in the presence of genes encoding iron acquisition systems and hypermucoid regulators between host types was identified using the Fisher's exact test implemented in Scoary as mentioned above.

All gene clusters produced by Panaroo were filtered to identify those associated with host type. Clusters encoding a single product were retained, while those annotated with differing functions or solely as a hypothetical protein were excluded. Gene clusters encoding both a hypothetical protein and known product were retained, with the annotation recorded as the known product, assuming the hypothetical protein also encoded it. Statistical significance was completed using a Wilcoxon rank-sum test in R to compare the distribution of count data between hosts [30]. The products of significant gene clusters were manually grouped into functional categories such as carbon sources, amino acid transport, DNA recombination, and plasmid association by searching each gene on UniProt, selecting the corresponding gene entry that matches the annotation provided by Prokka, and recording the biological function [31].

3. Results

3.1. Isolates and genome characteristics

A total of 128 KpI genomes (human, $n = 63$; bovine, $n = 65$) were used in this study. The number of contigs within each genome ranged from 22 to 309, with an N50 between 32,624 bp and 798,846 bp, and total genome size between 5.19 and 5.78 Mbp. The average coverage of each genome was between 15× and 113× [17]. Given the difference in isolation years between bovine and human isolates exceeding 10 years, our study emphasizes pangenome characteristics, presence of mobile genetic elements (MGEs), virulence traits, and metabolic pathways.

3.2. Large accessory genome and open pangenome in KpI

The pangenome of KpI consisted of 15,852 unique gene clusters. Core gene clusters (core genes and soft-core genes) comprised approximately 26 % of the pangenome, while accessory gene clusters (shell genes and cloud genes) accounted for approximately 74 %, highlighting the plastic and dynamic properties of KpI (Fig. 1A). The core genome of bovine KpI genomes was 37 % of the pangenome, while the core genome of human isolates was 33 % (Fig. 1A).

As new KpI genomes are analyzed, the number of unique gene clusters accumulates, indicating that the pangenome is open (Fig. 1B). Bovine KpI have a lower number of unique gene clusters ($n = 11{,}377$) and a larger core genome ($n = 3994$, ~35 % of bovine KpI pangenome) compared to human KpI which have approximately 12,896 unique gene

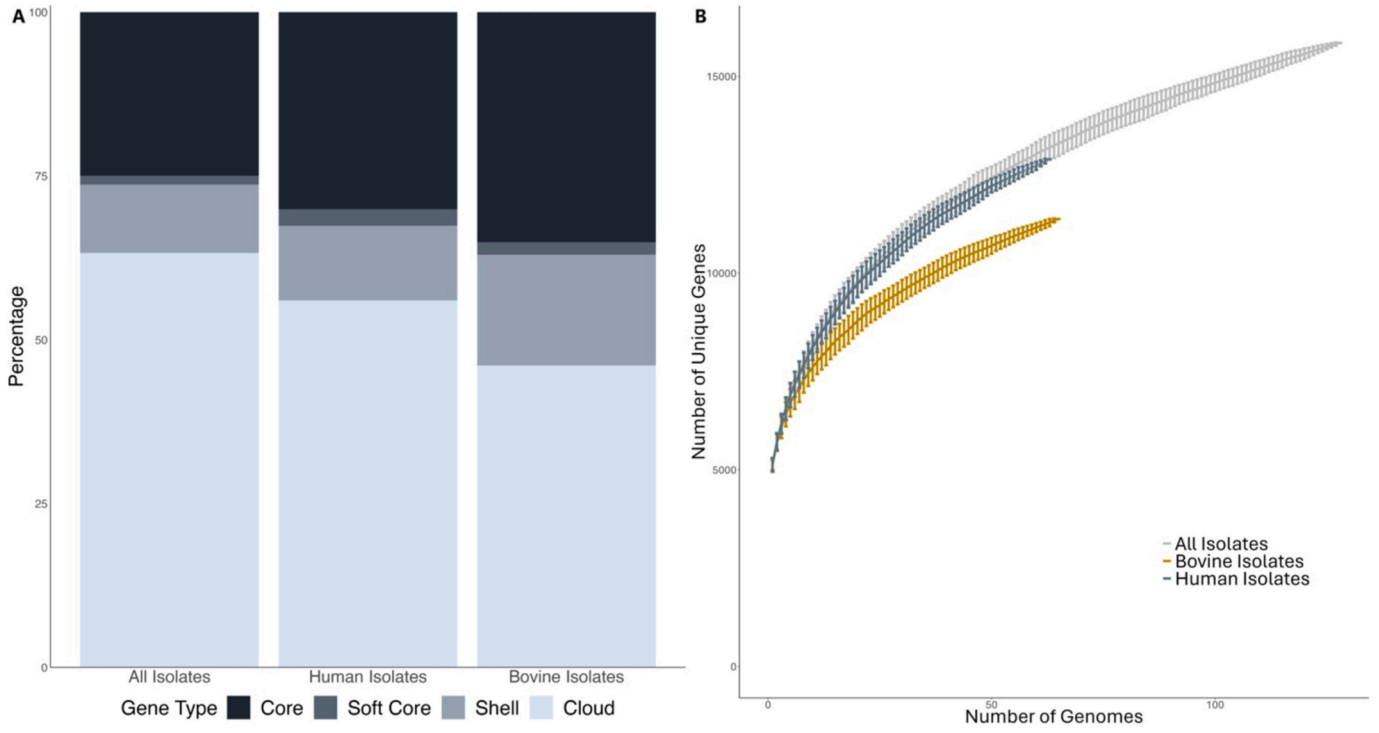


Fig. 1. Pangenome of *K. pneumoniae sensu stricto.*
(A) A stacked bar chart displays the relative abundance of the Kpl pangenome classified as core genes (present in 99 %–100 % of isolates), soft core genes (present in 95 %–98 % of isolates), shell genes (present in 15 %–94 % of isolates), and cloud genes (present in 1 % to 15 % of strains). (B) Gene accumulation curves display the number of unique genes per genome assembly and indicates an open pangenome.

clusters and a core genome size of 3879 genes (~30 % of human KpI pangenome). These findings suggest that bovine isolates may have lower genetic diversity than those isolated from humans.

3.3. Multi-locus sequence types and predicted serotypes

In total, 61 different STs of KpI were identified. STs are determined based on the allelic profile of seven housekeeping genes in *K. pneumoniae*, namely *gapA, infB, mdh, pgi, phoE, rpoB,* and *tonB* [32]. Among isolates from bovine hosts, 24 STs were identified, with ST2640 ($n = 12$), ST294 ($n = 9$), and ST107 ($n = 9$) being the most common. In comparison, 37 STs were reported in isolates from human hosts, with ST15 ($n = 5$), ST17 ($n = 4$), and ST45 ($n = 4$) being the most common. One ST, namely ST219, was shared among only one bovine isolate and one human isolate.

The K locus of *K. pneumoniae* contains several conserved genes including *galF, cpsACP, wzi, wza, wzb* and *wzc* located at the 5′, as well as gnd and *ugd* at the 3′ end, and encodes proteins involved in capsule synthesis and translocation [33]. Forty-seven different K-locus (KL) types with high confidence were identified among KpI genomes. Assemblies with low or no confidence of the KL type were classified as 'unknown KL' ($n = 26$). Most unknown KL types were from ST2640 ($n = 12$) and ST2650-ILV ($n = 6$) genomes—all of which were missing 3 KL51 genes, apart from one ST2640 strain which was missing 4 KL51 genes. Other STs with low or no KL confidence were ST1380 ($n = 1$, missing 5 genes from KL62), ST15 ($n = 1$, missing 10 genes from KL24), ST327-ILV ($n = 1$, missing 5 genes from KL143), ST3689 ($n = 1$, missing 2 genes from KL61), ST55-2LV ($n = 1$, missing 4 genes from KL124), ST5624 ($n = 2$, missing 4 genes from KL12), and ST569 ($n = 1$, missing 4 genes from KL143).

The O locus of *K. pneumoniae* contains *wzm* and *wzt* genes and encodes the membrane transporter complex [34]. Fourteen O-locus (OL) types were identified among isolates. Three isolates had low or no OL confidence, including ST1380 ($n = 1$; 6 genes missing from O3/O3a locus), ST2854 ($n = 1$, 2 genes missing from OL101 locus), and ST268 ($n = 1$; 1 gene missing from OL104 locus).

3.4. Bovine and human KpI core genomes are intermixed

An approximately-maximum-likelihood phylogenomic tree representing core genes of all isolates showed deep branching, indicating high diversity among all genomes. No clustering of genomes by host was observed, but rather isolates clustered by ST. For example, large clusters of bovine isolates pertaining to ST2640 & ST2640-1LV, ST107, and ST294 can be observed (Fig. 2).

3.5. Human and bovine KpI genomes contain the same plasmid replicons

A total of 28 unique plasmid replicons were identified among all KpI assemblies. There was no significant difference in the number of plasmid replicons found between bovine isolates (0–9) and human isolates (0−11) (Wilcoxon rank sum test, W = 2425, *p*-value = 0.07) (Fig. 3A). Eighteen plasmid replicons were found exclusively in either human and bovine STs, while the remaining 10 replicons were shared between the two groups. Exclusive replicons were less common, occuring in 1.5 % to 14 % of isolates, while shared replicons were more prevalent, occuring in 3 % to 95 % of isolates. Replicon type IncHI1B_1_pNDM-MAR (*p*-value <1E-10; OR = 19.4), RepA_1_pKPC-CAV1321 (p-value <0.05; OR = inf), IncHI2A_1 (p-value <0.05; OR = inf), and IncHI2_1 (p-value <0.05; OR = inf) were significantly associated with bovine KpI, while human KpI was significantly associated with IncFIB(K)_1_Kpn3 (p-value <0.05; OR = 0.17) and IncR_1 (p-value <0.05; OR = 9.8) (Fig. 3B). Due to the limitations with short-read sequencing data, we were unable to discern the contents of each plasmid. Our results indicate that host niche does not hinder plasmid acquisition, as common plasmid replicons are often found in genome assemblies of isolates from both bovine and human origin, this suggests that human and bovine KpI are capable of sharing a pool of MGEs.

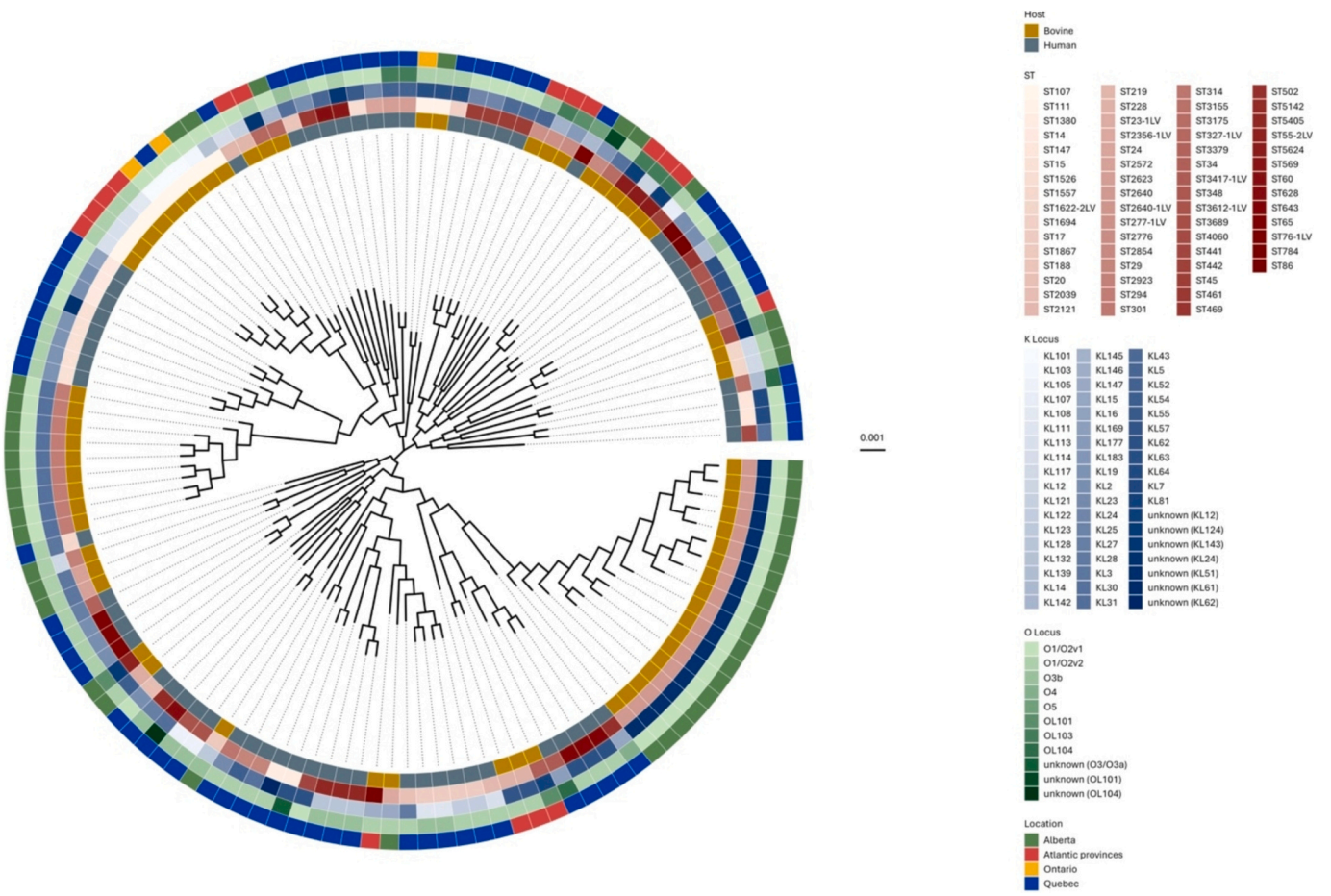


Fig. 2. Approximately-maximum-likelihood core phylogenomic tree of *K. pneumoniae sensu stricto* from bovine and human origin sequenced in this study. All genomes classified as *K. pneumoniae sensu stricto,* show deep branching and clustering by ST. External coloured boxes are used to indicate host-species, ST, K-locus, O-locus, and the location the sample that yielded an isolate was collected.

3.6. Genomic differences between bovine and human KpI

Given that iron acquisition systems are a main virulence factor studied in KpSC, we compared genes encoding iron acquisition systems between bovine and human KpI assemblies within our study. As identified by Kleborate, 18 KpI assemblies carried yersiniabactin, with 2 isolated from bovine hosts (p-value <0.001; OR = 11) (Fig. 4). Among the yersiniabactin lineages identified, ybt10; ICE_{Kp4} (n = 10) was the most common type, specifically among ST34 (n = 1), ST45 (n = 4), ST14 (n = 1), and ST15 (n = 4) (Fig. 4). Other yersiniabactin ICE elements included ybt12; ICE_{Kp10} (n = 1), ybt 14; ICE_{Kp5} (n = 2), ybt 15; ICE_{Kp11} (n = 1), ybt 2; ICE_{Kp1} (n = 1), ybt 26; ICE_{Kp18} (n = 1), and ybt 9; ICE_{Kp3} (n = 2). Seven human KpI assemblies contained genes encoding aerobactin (p-value <0.01; OR = inf), with 6 carrying *iuc* 1 lineage, and only 1 carrying *iuc* 2 lineage. Similarly, 7 human KpI assemblies carried salmochelin (p-value <0.01; OR = inf), with 5 carrying *iro* 1 lineage, 1 carrying the *iro* 2 lineage, and 1 carrying a truncated form of *iro* 3. One human isolate (ST2039) carried all four iron acquisition systems in addition to genes encoding colibactin, which is often co-mobilized on ICEKp10 (Fig. 4).

Regulators of a hypermucoid phenotype, *rmpADC* locus and possibly *rmpA2*, are often associated with hypervirulent infections in humans and therefore their presence was compared between human and bovine KpI assemblies [11]. Human KpI assemblies commonly carried capsule regulatory genes, *rmpADC* (n = 7; p-value <0.01; OR = inf) and *rmpA2* (n = 5; p-value <0.05; OR = inf), which included rmp 1 (n = 5), rmp 2 (n = 1), and rmp 3 lineages (n = 1), as well as *rmpA2*_6 lineage, respectively (Fig. 4). Bovine KpI assemblies were significantly associated with *fecABCDEIR* encoding the ferric citrate iron uptake system (p-value <0.001; OR = 19), with ~94 % of bovine KpI assemblies encoding it entirely (Fig. 4).

It was of interest if any other genes from bovine and human hosts were related to their respective environments. Among all gene clusters, 9687 encoded a hypothetical protein and 249 encoded multiple genes with multiple functions, and therefore were removed from the analysis. Gene clusters annotated as encoding both hypothetical protein as well as a known product were retained (n = 457) and the annotation was counted as the known product, assuming the hypothetical protein also encoded this product. The remaining gene clusters (n = 5469) encoded only one product. Since different gene clusters can encode the same product, we generated a non-redundant list of 3658 gene products from all filtered gene clusters. We then examined significance in the number of gene clusters encoding products from this list between host types.

Of the 3658 gene annotations screened for host association, 392 were deemed significant (p-value <0.01) and manually categorized into relevant groupings (Supplementary file S1). Bovine KpI genomes contained a significantly higher number of gene clusters associated with histidine (*hisQMP*) (p-value <0.001, W = 3150.0) and arginine transport proteins (*astABCDE* and *argT*) (p-value <0.001, 3024 ≤ W ≤ 3085), and a signficantly lower number of gene clusters encoding glutathione transport (*gsiABCD*) (p-value <0.001, 1309.0 ≤ W ≤ 1490.5) in comparison to human KpI genomes (Fig. 5A). Moreover, bovine assemblies contained more gene clusters involved in the use of galactonate (*lgoRDT*) (p-value <0.001, 2691 ≤ W ≤ 2921.5), galactose (*dgoAKD*, *galP*, *gci*) (p-value <0.001, 2457 ≤ W ≤ 2583), and lactose (*lacZ*) (p-value <0.001, W = 3339.5), while human assemblies contained more gene clusters

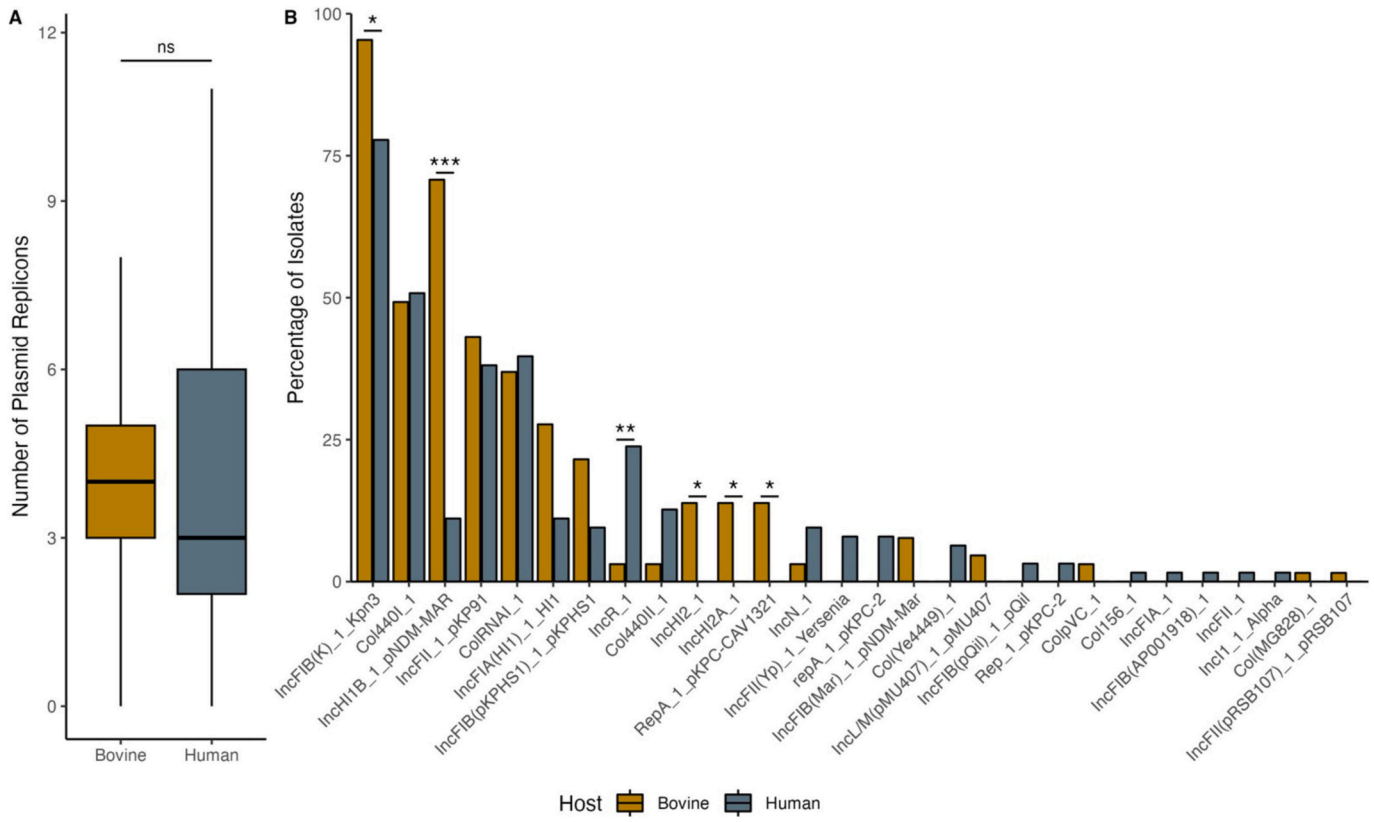


Fig. 3. Plasmid replicons identified in *K. pneumoniae sensu stricto* STs from bovine and human hosts.
Assemblies in this study were examined for known plasmid replicons using the PlasmidFinder database. (A) The number of plasmid replicons found in genome assemblies is not statistically significant between human and bovine hosts. (B) Some plasmid replicons show significant association with host type, including IncHI1B_1_pNDM-MAR, RepA_1_pKPC-CAV1321, IncHI2A_1, and IncHI2_1 with bovine KpI genomes and IncFIB(K)_1_Kpn3 and IncR_1 with human KpI assemblies. Statistical significance was completed using a Benjamini-Hochberg adjusted Fisher's exact *t*-test (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$).

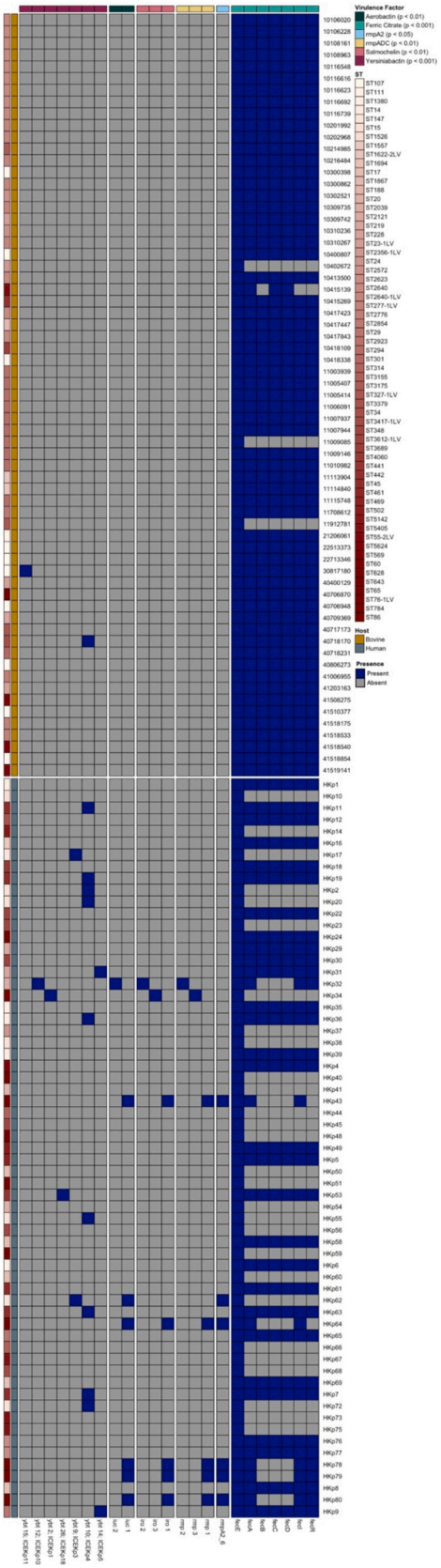
involved in utilization of fructose (*fruA*) (p-value <0.05, W = 1667), gluconate (p-value <0.01, W = 1418.5), sucrose (*scrY*) (p-value <0.01, W = 1082.5), inositol (*mshA*, *iolEGSU*) (p-value <0.05, $1410.0 \leq W \leq 1542.5$), and glycerol (*dhaBT*, *glpF*) (p-value <0.01, $1410 \leq W \leq 1480.5$) (Fig. 5B). Notably, some bovine assemblies contained up to 5 gene clusters encoding a beta-galactosidase (*lacZ*), while human assemblies contained at most 3.

4. Discussion

Klebsiella is thought of as an unrecognized zoonotic pathogen, capable of spreading directly from animals to humans or indirectly through bovine products such as raw milk [13]. In this study, we identified high genomic similarity between KpI assemblies from cases of bovine mastitis and human blood infections. Overall, no obvious genetic barriers to transmission between cows and humans were identified, which contrasts previous reports of other mastitis pathogens such as *Staphylococcus aureus* [35]. Rather, subtle genomic differences were identified in bovine KpI which likely reflect environmental adaptation.

The frequency of overlapping STs between bovine and human isolates in our study was low, with only ST219 being shared between one bovine and one human isolate, implying that transmission between bovine and humans may be limited as others have suggested [36,37]. However, limited ST overlap between hosts within our study may be a result of small sample size, as numerous bovine STs from our study have been identified in humans elsewhere, including predominant bovine STs, ST107, ST294, and ST2640 [38–40]. These dominant STs in bovine isolates likely resulted in the lower diversity observed compared to human isolates, since samples from the same lineage often have high genomic similarity within core and common genes [5]. ST2640/ST2640-1LV and ST294 were collected from farms in Alberta which may explain the low variance recorded, while ST107 was isolated from all four provinces sampled. Other studies have reported ST107 as the predominant ST from clinical mastitis, bulk milk, and environmental sources [41,42]. Although ST107 is common among bovine hosts and the dairy farm environment, it is not restricted to it, as it has been identified in other hosts such as captive alligators, wild tapirs, and humans [38,39,43,44].

Subtle genomic differences between host types were observed in our study that may provide insight into the development of novel host-specific therapeutic options, improving improve health outcomes in humans and animals. The most evident genomic differences between bovine and human KpI were the number of clusters associated with LacZ β-galactosidase and ferric citrate uptake. The *fec* and *lac* operons are known to be co-located together on various plasmids, including plasmid pKPN3, the replicon of which was identified in most bovine isolates in this study [1]. Indeed, within this study additional copies of *fec* genes and *lacZ* were found to be present within both bovine and human KpI assemblies often adjacent to transposases, suggesting these may be mobile [1]. While the entire *fec* operon was abundant within bovine isolates in our study, only gene clusters associated with *lacZ,* encoding an orphan β-galactosidase, which cleaves lactose into glucose and galactose, was in higher abundance, as opposed to the entire *lac* operon (*lacZYA*) reported in other studies [1]. Given the high abundance of lactose within milk, bovine *K. pneumoniae* may have evolved higher copy numbers of *lacZ* to allow it to efficiently utilize lactose as a carbon source in the mammary gland. Additionally, bovine KpI isolates show a higher abundance of genes associated with galactose and galactonate, compounds derived from lactose metabolism [45], which may also be associated growth advantages for KpI during bovine mastitis.



(caption on next column)

Fig. 4. Heatmap of Virulence Factors in Bovine and Human KpI Genomes. Genes encoding the ferric citrate iron uptake system (*fecABCDEIR*) were highly prevalent in bovine assemblies, and less prevalent in human KpI assemblies. Conversely, human KpI assemblies contained other iron acquisition systems like salmochelin, aerobactin, and yersiniabactin. Human KpI assemblies also contained more hypermucoid regulatory genes, *rmpADC* locus and *rmpA2*. To test if the presence of certain virulence factors was statistically different between host type, statistical significance was completed using a Benjamini-Hochberg adjusted Fisher's exact t-test and *p*-values for each test are recorded in the Virulence Factor legend of the plot.

In addition to lactose, citrate is present in milk and can bind iron with high affinity [46]. The concentration of citrate in bovine milk ranges between 11.2 and 15.1 mmol/L, varies depending on the stage of lactation, and is typically higher in concentration than other iron chelating agents [47,48]. Proteins encoded by the *fec* locus bind and translocate iron from citrate, providing a source of iron for bacteria in milk. Ferric citrate uptake has been linked to the pathogenicity of mammary pathogenic *E. coli* (MPEC), and deletion of the *fec* locus in MPEC resulted in the inability for mastitis to progress [49]. The higher prevelence of *fecIRABCDE* in bovine KpI genomes compared to human suggests that its expression may be an important virulence factor for bovine KpI. However, further *in vitro* and *in vivo* studies are needed to confirm the role of the *fec* locus in the virulence of bovine KpI.

A higher abundance of genes associated with histidine and arginine use was found in bovine KpI isolates which provides further evidence of KpI host adaptation, that has yet to be reported. Histidine and arginine can be supplemented to dairy cows to increase milk protein yield and milk fat production, respectively, and thus both can be found within raw milk, providing a source of nitrogen that may be used by KpI within the mammary gland [50,51].

Our study focused on bovine KpI from mastitis, however future research should aim to compare commensal KpI from dairy environments with KpI from mastitis to better understand host specificity and mechanisms driving pathogenicity. Moreover, draft genomes used in this study, while assesed for quality using contig count, coverage,and N50, may still be incomplete due to gaps, sequencing errors, and/or missassblies [52]. Additionally, the time gap between isolating human and bovine KpI strains may have driven differences observed in virulence, and metabolic processes and thus results should be interpreted with caution. Despite these limitations, our findings provide valuable insights into the genomic differences and similarities between human and animal KpI isolates, advancing our understanding of the pathogenic dynamic across all hosts, and strengthening One Health strategies to mitigate the spread of KpSC across sectors.

Funding sources

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC). The research was funded by an NSERC Alliance International Catalyst grant to JR, SF, and MS.

CRediT authorship contribution statement

Bridget O'Brien: Writing – original draft, Methodology, Investigation, Formal analysis. **Alla Yushchenko:** Writing – review & editing, Methodology, Formal analysis. **Jinha Suh:** Writing – review & editing, Investigation, Formal analysis. **Dongyun Jung:** Writing – review & editing, Resources, Methodology, Investigation. **Zhangbin Cai:** Writing – review & editing, Resources, Methodology, Investigation. **Ngoc Sang Nguyen:** Writing – review & editing, Resources, Methodology, Investigation. **Makeda Semret:** Writing – review & editing, Supervision, Resources, Funding acquisition. **Simon Dufour:** Writing – review & editing, Resources, Funding acquisition. **Séamus Fanning:** Writing – review & editing, Supervision, Resources, Investigation, Funding acquisition, Conceptualization. **Jennifer Ronholm:** Writing – review &

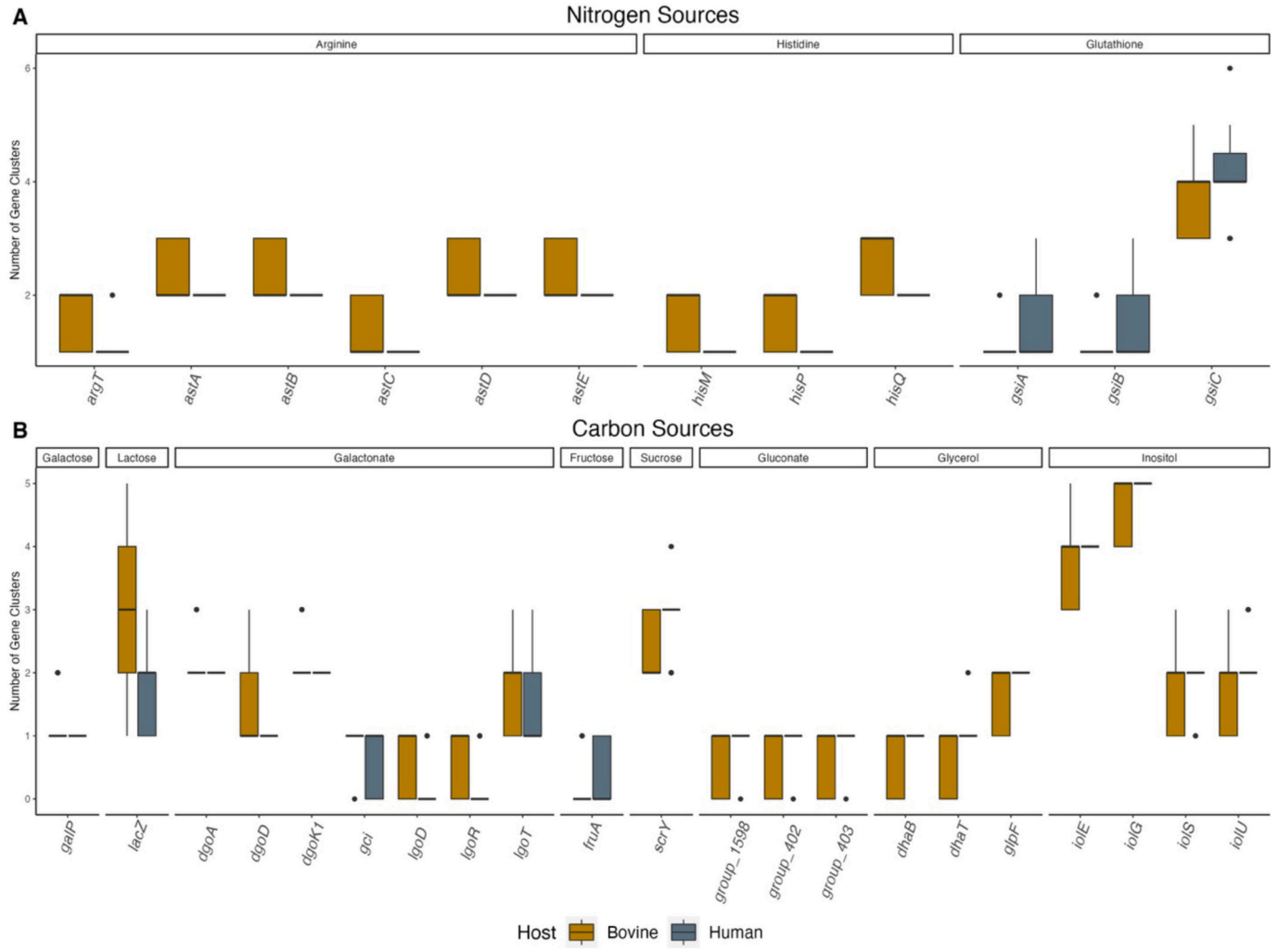


Fig. 5. Number of Gene Clusters Encoding Nitrogen and Carbon Utilization in Bovine and Human KpI.
(A) Bovine KpI genomes contained a significantly higher number of clusters associated with the uptake of arginine and histidine, while human KpI genomes contained a greater amount clusters associated with use of glutathione. (B) The abundance of gene clusters associated with utilization of carbon varied between bovine and human KpI depending on the carbon source. Bovine KpI genomes contained significantly more clusters associated with the use of lactose, galactonate, and galactose. Human KpI genomes contained more gene clusters associated with the utilization of fructose, gluconate, sucrose, inositol, and glycerol. X-axis labels indicate the common gene name among gene clusters. Specific gene clusters associated with each function are listed the supplementary data. Unique gene clusters *group_402*, *group_1598*, and *group_403* are associated with gluconate 2-dehydrogenase cytochrome *c* subunit, gluconate 2-dehydrogenase flavoprotein, and gluconate 2-dehydrogenase subunit 3, respectively. All box plots illustrate data from all KpI genomes within our study.

editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.onehlt.2025.100970.

Data availability

Data will be made available on request.

References

[1] K.E. Holt, H. Wertheim, R.N. Zadoks, S. Baker, C.A. Whitehouse, D. Dance, A. Jenney, T.R. Connor, L.Y. Hsu, J. Severin, S. Brisse, H. Cao, J. Wilksch, C. Gorrie, M.B. Schultz, D.J. Edwards, K.V. Nguyen, T.V. Nguyen, T.T. Dao, M. Mensink, V. L. Minh, N.T.K. Nhu, C. Schultsz, K. Kuntaman, P.N. Newton, C.E. Moore, R. A. Strugnell, N.R. Thomson, Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella pneumoniae, an urgent threat to public health, Proc. Natl. Acad. Sci. 112 (2015) E3574–E3581, https://doi.org/10.1073/pnas.1501049112.

[2] N. Raffelsberger, M.A.K. Hetland, K. Svendsen, L. Småbrekke, I.H. Löhr, L.L. E. Andreassen, S. Brisse, K.E. Holt, A. Sundsfjord, Ø. Samuelsen, K. Gravningen, Gastrointestinal carriage of Klebsiella pneumoniae in a general adult population: a cross-sectional study of risk factors and bacterial genomic diversity, Gut Microbes 13 (2024) 1939599, https://doi.org/10.1080/19490976.2021.1939599.

[3] C.L. Gorrie, M. Mirceta, R.R. Wick, D.J. Edwards, N.R. Thomson, R.A. Strugnell, N. F. Pratt, J.S. Garlick, K.M. Watson, D.V. Pilcher, S.A. McGloughlin, D.W. Spelman, A.W.J. Jenney, K.E. Holt, Gastrointestinal carriage is a major reservoir of Klebsiella pneumoniae infection in intensive care patients, Clin. Infect. Dis. 65 (2017) 208–215, https://doi.org/10.1093/cid/cix270.

[4] B.-T. Huynh, V. Passet, A. Rakotondrasoa, T. Diallo, A. Kerleguer, M. Hennart, A. D. Lauzanne, P. Herindrainy, A. Seck, R. Bercion, L. Borand, M. Pardos de la Gandara, E. Delarocque-Astagneau, D. Guillemot, M. Vray, B. Garin, J.-M. Collard, C. Rodrigues, S. Brisse, Klebsiella pneumoniae carriage in low-income countries: antimicrobial resistance, genomic diversity and risk factors, Gut Microbes 11 (2020) 1287–1299, https://doi.org/10.1080/19490976.2020.1748257.

[5] C.L. Gorrie, M. Mirčeta, R.R. Wick, L.M. Judd, M.M.C. Lam, R. Gomi, I.J. Abbott, N. R. Thomson, R.A. Strugnell, N.F. Pratt, J.S. Garlick, K.M. Watson, P.C. Hunter, D. V. Pilcher, S.A. McGloughlin, D.W. Spelman, K.L. Wyres, A.W.J. Jenney, K.E. Holt, Genomic dissection of Klebsiella pneumoniae infections in hospital patients reveals

insights into an opportunistic pathogen, Nat. Commun. 13 (2022) 3017, https://doi.org/10.1038/s41467-022-30717-6.
[6] Y.T. Gröhn, D.J. Wilson, R.N. González, J.A. Hertl, H. Schulte, G. Bennett, Y. H. Schukken, Effect of pathogen-specific clinical mastitis on milk yield in dairy cows, J. Dairy Sci. 87 (2004) 3358–3374, https://doi.org/10.3168/jds.S0022-0302(04)73472-4.
[7] R.N. Zadoks, H.M. Griffiths, M.A. Munoz, C. Ahlstrom, G.J. Bennett, E. Thomas, Y. H. Schukken, Sources of Klebsiella and Raoultella species on dairy farms: be careful where you walk, J. Dairy Sci. 94 (2011) 1045–1051, https://doi.org/10.3168/jds.2010-3603.
[8] J. Massé, S. Dufour, M. Archambault, Characterization of Klebsiella isolates obtained from clinical mastitis cases in dairy cattle, J. Dairy Sci. 103 (2020) 3392–3400, https://doi.org/10.3168/jds.2019-17324.
[9] M.K. Paczosa, J. Mecsas, Klebsiella pneumoniae: going on the offense with a strong defense, Microbiol. Mol. Biol. Rev. 80 (2016) 629–661, https://doi.org/10.1128/MMBR.00078-15.
[10] J.B. Cheng, J.Q. Wang, D.P. Bu, G.L. Liu, C.G. Zhang, H.Y. Wei, L.Y. Zhou, J. Z. Wang, Factors affecting the lactoferrin concentration in bovine milk, J. Dairy Sci. 91 (2008) 970–976, https://doi.org/10.3168/jds.2007-0689.
[11] J.C. Catalán-Nájera, U. Garza-Ramos, H. Barrios-Camacho, Hypervirulence and hypermucoviscosity: two different but complementary Klebsiella spp. phenotypes? Virulence 8 (2017) 1111–1123, https://doi.org/10.1080/21505594.2017.1317412.
[12] K.L. Wyres, K.E. Holt, Klebsiella pneumoniae as a key trafficker of drug resistance genes from environmental to clinically important bacteria, Curr. Opin. Microbiol. 45 (2018) 131–139, https://doi.org/10.1016/j.mib.2018.04.004.
[13] Y. Hu, J. Anes, S. Devineau, S. Fanning, Klebsiella pneumoniae: prevalence, reservoirs, antimicrobial resistance, pathogenicity, and infection: a hitherto unrecognized zoonotic bacterium, Foodborne Pathog. Dis. 18 (2021) 63–84, https://doi.org/10.1089/fpd.2020.2847.
[14] G.S. Davis, L.B. Price, Recent research examining links among Klebsiella pneumoniae from food, food animals, and human extraintestinal infections, Curr. Environ. Health Rpt. 3 (2016) 128–135, https://doi.org/10.1007/s40572-016-0089-9.
[15] Z. Zheng, P.J. Gorden, X. Xia, Y. Zheng, G. Li, Whole-genome analysis of Klebsiella pneumoniae from bovine mastitis milk in the U.S, Environ. Microbiol. 24 (2022) 1183–1199, https://doi.org/10.1111/1462-2920.15721.
[16] S. Dufour, J. Labrie, M. Jacques, The mastitis pathogens culture collection, Microbiol. Resour. Announc. 8 (2019), https://doi.org/10.1128/MRA.00133-19 e00133-19.
[17] B. O'Brien, A. Yushchenko, J. Suh, D. Jung, Z. Cai, N.S. Nguyen, M. Semret, S. Dufour, J. Ronholm, Draft genome sequences of 148 Klebsiella pneumoniae species complex members from bovine and human hosts, Microbiol. Resour. Announc. 0 (2024), https://doi.org/10.1128/mra.00132-24 e00132-24.
[18] M.M.C. Lam, R.R. Wick, S.C. Watts, L.T. Cerdeira, K.L. Wyres, K.E. Holt, A genomic surveillance framework and genotyping tool for Klebsiella pneumoniae and its related species complex, Nat. Commun. 12 (2021) 4188, https://doi.org/10.1038/s41467-021-24448-3.
[19] R.R. Wick, E. Heinz, K.E. Holt, K.L. Wyres, Kaptive web: user-friendly capsule and lipopolysaccharide serotype prediction for Klebsiella genomes, J. Clin. Microbiol. 56 (2018), https://doi.org/10.1128/jcm.00197-18.
[20] G. Tonkin-Hill, N. MacAlasdair, C. Ruis, A. Weimann, G. Horesh, J.A. Lees, R. A. Gladstone, S. Lo, C. Beaudoin, R.A. Floto, S.D.W. Frost, J. Corander, S. D. Bentley, J. Parkhill, Producing polished prokaryotic pangenomes with the Panaroo pipeline, Genome Biol. 21 (2020) 180, https://doi.org/10.1186/s13059-020-02090-4.
[21] T. Seemann, Prokka: rapid prokaryotic genome annotation, Bioinformatics 30 (2014) 2068–2069, https://doi.org/10.1093/bioinformatics/btu153.
[22] K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, Nucleic Acids Res. 30 (2002) 3059–3066, https://doi.org/10.1093/nar/gkf436.
[23] J. Oksanen, vegan: Community Ecology Package, R package version 2.6-4, https://CRAN.R-project.org/package=vegan, 2024.
[24] M.N. Price, P.S. Dehal, A.P. Arkin, FastTree 2 – approximately maximum-likelihood trees for large alignments, PLoS One 5 (2010) e9490, https://doi.org/10.1371/journal.pone.0009490.
[25] G. Yu, D.K. Smith, H. Zhu, Y. Guan, T.T.-Y. Lam, Ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data, Methods Ecol. Evol. 8 (2017) 28–36, https://doi.org/10.1111/2041-210X.12628.
[26] A. Carattoli, E. Zankari, A. García-Fernández, M. Voldby Larsen, O. Lund, L. Villa, F. Møller Aarestrup, H. Hasman, In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing, Antimicrob. Agents Chemother. 58 (2014) 3895–3903, https://doi.org/10.1128/AAC.02412-14.
[27] T. Seemann, Abricate. https://github.com/tseemann/abricate, 2024.
[28] S. Arredondo-Alonso, M.R.C. Rogers, J.C. Braat, T.D. Verschuuren, J. Top, J. Corander, R.J.L. Willems, A.C. Schürch, Mlplasmids: a user-friendly tool to predict plasmid- and chromosome-derived sequences for single species, Microbial. Genom. 4 (2018) e000224, https://doi.org/10.1099/mgen.0.000224.
[29] O. Brynildsrud, J. Bohlin, L. Scheffer, V. Eldholm, Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary, Genome Biol. 17 (2016) 238, https://doi.org/10.1186/s13059-016-1108-8.
[30] R Core Team, R: A Language and Environment for Statistical Computing. https://www.R-project.org/, 2023.
[31] The UniProt Consortium, UniProt: the universal protein knowledgebase in 2023, Nucleic Acids Res. 51 (2023) D523–D531, https://doi.org/10.1093/nar/gkac1052.
[32] L. Diancourt, V. Passet, J. Verhoef, P.A.D. Grimont, S. Brisse, Multilocus sequence typing of Klebsiella pneumoniae nosocomial isolates, J. Clin. Microbiol. 43 (2005) 4178, https://doi.org/10.1128/JCM.43.8.4178-4182.2005.
[33] Y.-J. Pan, T.-L. Lin, C.-T. Chen, Y.-Y. Chen, P.-F. Hsieh, C.-R. Hsu, M.-C. Wu, J.-T. Wang, Genetic analysis of capsular polysaccharide synthesis gene clusters in 79 capsular types of Klebsiella spp, Sci. Rep. 5 (2015) 15573, https://doi.org/10.1038/srep15573.
[34] L. Cuthbertson, M.S. Kimber, C. Whitfield, Substrate binding by a bacterial ABC transporter involved in polysaccharide export, Proc. Natl. Acad. Sci. USA 104 (2007) 19529–19534, https://doi.org/10.1073/pnas.0705709104.
[35] S. Park, D. Jung, B. O'Brien, J. Ruffini, F. Dussault, A. Dube-Duquette, É. Demontier, J.-F. Lucier, F. Malouin, S. Dufour, J. Ronholm, Comparative genomic analysis of *Staphylococcus aureus* isolates associated with either bovine intramammary infections or human infections demonstrates the importance of restriction-modification systems in host adaptation, Microbial. Genom. 8 (2024) (2022) 000779, https://doi.org/10.1099/mgen.0.000779.
[36] H.A. Thorpe, R. Booton, T. Kallonen, M.J. Gibbon, N. Couto, V. Passet, S. López-Fernández, C. Rodrigues, L. Matthews, S. Mitchell, R. Reeve, S. David, C. Merla, M. Corbella, C. Ferrari, F. Comandatore, P. Marone, S. Brisse, D. Sassera, J. Corander, E.J. Feil, A large-scale genomic snapshot of Klebsiella spp. isolates in northern Italy reveals limited transmission between clinical and non-clinical settings, Nat. Microbiol. 7 (2022) 2054–2067, https://doi.org/10.1038/s41564-022-01263-0.
[37] A. Dereeper, G. Gruel, M. Pot, D. Couvin, E. Barbier, S. Bastian, J.-C. Bambou, M. Gelu-Simeon, S. Ferdinand, S. Guyomard-Rabenirina, V. Passet, F. Martino, P. Piveteau, Y. Reynaud, C. Rodrigues, P.-M. Roger, X. Roy, A. Talarmin, B. Tressieres, M. Valette, S. Brisse, S. Breurec, Limited transmission of Klebsiella pneumoniae among humans, animals, and the environment in a Caribbean Island, Guadeloupe (French West Indies), Microbiol. Spectr. 10 (2024), https://doi.org/10.1128/spectrum.01242-22 e01242-22.
[38] X. Wang, H. Li, C. Zhao, H. Chen, J. Liu, Z. Wang, Q. Wang, Y. Zhang, W. He, F. Zhang, H. Wang, Novel NDM-9 metallo-β-lactamase identified from a ST107 *Klebsiella pneumoniae* strain isolated in China, Int. J. Antimicrob. Agents 44 (2014) 90–91, https://doi.org/10.1016/j.ijantimicag.2014.04.010.
[39] L. Fu, M. Huang, X. Zhang, X. Yang, Y. Liu, L. Zhang, Z. Zhang, G. Wang, Y. Zhou, Frequency of virulence factors in high biofilm formation *bla*KPC-2 producing *Klebsiella pneumoniae* strains from hospitals, Microb. Pathog. 116 (2018) 168–172, https://doi.org/10.1016/j.micpath.2018.01.030.
[40] Y.-T. Huang, Y.-W. Kuo, N.-Y. Lee, N. Tien, C.-H. Liao, L.-J. Teng, W.-C. Ko, P.-R. Hsueh, Evaluating NG-test CARBA 5 multiplex Immunochromatographic and Cepheid Xpert CARBA-R assays among Carbapenem-resistant Enterobacterales isolates associated with bloodstream infection, Microbiol. Spectr. 10 (2024), https://doi.org/10.1128/spectrum.01728-21 e01728-21.
[41] J. Yang, Y. Xiong, H.W. Barkema, X. Tong, Y. Lin, Z. Deng, J.P. Kastelic, D. B. Nobrega, Y. Wang, B. Han, J. Gao, Comparative genomic analyses of Klebsiella pneumoniae K57 capsule serotypes isolated from bovine mastitis in China, J. Dairy Sci. (2023), https://doi.org/10.3168/jds.2023-23721.
[42] S. Song, W. He, D. Yang, M. Benmouffok, Y. Wang, J. Li, C. Sun, X. Song, S. Ma, C. Cai, S. Ding, C. Wu, Z. Shen, Y. Wang, Molecular epidemiology of Klebsiella pneumoniae from clinical bovine mastitis in northern area of China, 2018–2019, Engineering 10 (2022) 146–154, https://doi.org/10.1016/j.eng.2021.01.015.
[43] A.T.H.I. de Sousa, M.T.S. Costa, S.L. Cândido, H. Makino, T.O. Morgado, L.A. D. Pavelegini, E.M. Colodel, L. Nakazato, V. Dutra, Determination of multidrug-resistant populations and molecular characterization of complex Klebsiella spp. in wild animals by multilocus sequence typing, vet, World 15 (2022) 1691–1698, https://doi.org/10.14202/vetworld.2022.1691-1698.
[44] Y. Yang, C.H. Higgins, I. Rehman, K.N. Galvao, I.L. Brito, M.L. Bicalho, J. Song, H. Wang, R.C. Bicalho, Genomic diversity, virulence, and antimicrobial resistance of Klebsiella pneumoniae strains from cows and humans, Appl. Environ. Microbiol. 85 (2019), https://doi.org/10.1128/AEM.02654-18 e02654-18.
[45] D.K. Grange, Galactosemia, in: L.R. Johnson (Ed.), Encyclopedia of Gastroenterology, Elsevier, New York, 2004, pp. 83–87, https://doi.org/10.1016/B0-12-386860-2/00773-5.
[46] B. Lönnerdal, C.L. Keen, L.S. Hurley, Iron, copper, zinc, and manganese in milk, Annu. Rev. Nutr. 1 (1981) 149–174, https://doi.org/10.1146/annurev.nu.01.070181.001053.
[47] R.P. Glahn, C. Lai, J. Hsu, J.F. Thompson, M. Guo, D.R. Van Campen, Decreased citrate improves Iron availability from infant formula: application of an in vitro digestion/Caco-2 cell culture Model1, J. Nutr. 128 (1998) 257–264, https://doi.org/10.1093/jn/128.2.257.
[48] P. Hyvönen, T. Haarahiltunen, T. Lehtolainen, J. Heikkinen, R. Isomäki, S. Pyörälä, Concentrations of bovine lactoferrin and citrate in milk during experimental endotoxin mastitis in early- versus late-lactating dairy cows, J. Dairy Res. 77 (2010) 474–480, https://doi.org/10.1017/S0022029910000579.
[49] S.E. Blum, R.J. Goldstone, J.P.R. Connolly, M. Répérant-Ferter, P. Germon, N. F. Inglis, O. Krifucks, S. Mathur, E. Manson, K. Mclean, P. Rainard, A.J. Roe, G. Leitner, D.G.E. Smith, Postgenomics characterization of an essential genetic determinant of mammary pathogenic *Escherichia coli*, mBio 9 (2018), https://doi.org/10.1128/mbio.00423-18.

[50] H. Lapierre, G.E. Lobley, D.R. Ouellet, Histidine optimal supply in dairy cows through determination of a threshold efficiency, J. Dairy Sci. 104 (2021) 1759–1776, https://doi.org/10.3168/jds.2020-19205.
[51] L. Ding, Y. Shen, M. Jawad, T. Wu, S.K. Maloney, M. Wang, N. Chen, D. Blache, Effect of arginine supplementation on the production of milk fat in dairy cows, J. Dairy Sci. 105 (2022) 8115–8129, https://doi.org/10.3168/jds.2021-21312.
[52] N. Ricker, H. Qian, R.R. Fulthorpe, The limitations of draft assemblies for understanding prokaryotic adaptation and evolution, Genomics 100 (2012) 167–175, https://doi.org/10.1016/j.ygeno.2012.06.009.