Title

WGS of intrauterine E. coli from cows with early postpartum uterine infection reveals a non-uterine specific genotype and virulence factors.

Permalink

Journal

Authors

Garzon, Adriana

Basbas, Carl

Schlesener, Cory

et al.

Publication Date

DOI

Peer reviewed

WGS of intrauterine *E. coli* from cows with early postpartum uterine infection reveals a non-uterine specific genotype and virulence factors

Adriana Garzon,[1] Carl Basbas,[1] Cory Schlesener,[1,2] Noelia Silva-del-Rio,[1,3] Betsy M. Karle,[4] Fabio S. Lima,[1] Bart C. Weimer,[1,2] Richard V. Pereira[1]

AUTHOR AFFILIATIONS See affiliation list on p. 14.

ABSTRACT *Escherichia coli* has been attributed to playing a major role in a cascade of events that affect the prevalence and severity of uterine disease in cattle. The objectives of this project were to (i) define the association between the prevalence of specific antimicrobial resistance and virulence factor genes in *E. coli* with the clinical status related to uterine infection, (ii) identify the genetic relationship between *E. coli* isolates from cows with diarrhea, with mastitis, and with and without metritis, and (iii) determine the association between the phenotypic and genotypic antimicrobial resistance identified on the *E. coli* isolated from postpartum cattle. Bacterial isolates ($n = 148$) were obtained from a larger cross-sectional study. Cows were categorized into one of three clinical groups before enrollment: metritis, cows with purulent discharge, and control cows. For genomic comparison, public genomes ($n = 130$) from cows with diarrhea, mastitis, and metritis were included in a genome-wide association study, to evaluate differences between the drug classes or the virulence factor category among clinical groups. A distinct *E. coli* genotype associated with metritis could not be identified. Instead, a high genetic diversity among the isolates from uterine sources was present. A virulence factor previously associated with metritis (*fim*H) using PCR was not associated with metritis. There was moderate accuracy for whole-genome sequencing to predict phenotypic resistance, which varied depending on the antimicrobial tested. Findings from this study contradict the traditional pathotype classification and the unique intrauterine *E. coli* genotype associated with metritis in dairy cows.

IMPORTANCE Metritis is a common infectious disease in dairy cattle and the second most common reason for treating a cow with antimicrobials. The pathophysiology of the disease is complex and is not completely understood. Specific endometrial pathogenic *Escherichia coli* have been reported to be adapted to the endometrium and sometimes lead to uterine disease. Unfortunately, the specific genomic details of the endometrial-adapted isolates have not been investigated using enough genomes to represent the genomic diversity of this organism to identify specific virulence genes that are consistently associated with disease development and severity. Results from this study provide key microbial ecological advances by elucidating and challenging accepted concepts for the role of Intrauterine *E. coli* in metritis in dairy cattle, especially contradicting the existence of a unique intrauterine *E. coli* genotype associated with metritis in dairy cows, which was not found in our study.

KEYWORDS whole-genome sequencing, cattle, mutation, antibiotic resistance, uterine infection

Metritis is a common infectious disease in dairy cattle and the second most common reason for treating a cow with antibiotics (1). Metritis typically occurs within 21 days postpartum, characterized by an enlarged uterus and fetid, watery red-brown uterine discharge with systemic signs of illness (2). Metritis also has an important effect on herd profitability (3). The cost of a metritis case in cows treated with ampicillin and with ceftiofur has been estimated to average $344 and $410, respectively (4). However, the clinical and reproductive beneficial influence of antibiotic treatment has yielded inconsistent results (5). The variability in treatment success may be explained in part by the lack of a gold standard for disease diagnosis, the self-cure rate associated with the disease, and the variability of including control groups that hinder the comparison among studies.

The pathophysiology of the disease is complex with limited studies that generate information to allow causal inferences about the disease (6). Part of this challenge includes a lack of consensus on the criteria and terminology used for clinical disease diagnosis in research studies, as recently demonstrated in a scoping review (ScR) (7). Although inconsistencies in the clinical definition of metritis exist, the aforementioned ScR identified that most studies defined metritis as a fetid, red-brown watery discharge with or without systemic signs of disease based on the work of Sheldon et al. (2).

The reproductive tract in cows harbors a diverse microbiome (8–10), and recent advancements in next-generation sequencing have demonstrated that the uterus has an established microbiome even before parturition and does not exclusively result from microbial exposure during calving (11–13). Several bacteria have been isolated from metritic cows and are suspected as causal agents of metritis, including *Escherichia coli*, *Trueperella pyogenes*, *Fusobacterium necrophorum*, and *Prevotella* spp. (14, 15). Descriptive studies using 16S rRNA sequencing revealed a higher abundance of *Bacteroides*, *Porphyromonas*, and *Fusobacterium* species in cows with metritis when compared with healthy cows, while *Escherichia coli* bacteria were found to be more abundant in healthy animals (11, 16, 17). *In vivo* models of uterine infections have been successfully developed using pathogenic *T. pyogenes* alone (18), in combination with *E. coli* (19) or as a pathogenic cocktail of *T. pyogenes*, *E. coli*, and *F. necrophorum* (20), yet the disease causality has not been demonstrated with other bacterial species.

Based on current data for *Escherichia coli* in metritic cows, these bacteria have been proposed to play a major role in a cascade of events that affect the prevalence, severity, and persistence of uterine disease in cattle (6). Specific endometrial pathogenic *Escherichia coli* (EnPEC) have been described as adapted to the endometrium, allowing them to colonize this environment and develop uterine disease (21), as well as increasing the susceptibility of the endometrium to infection with *T. pyogenes* (22).

Most of the data for intrauterine *E. coli* from cows with metritis are either based on culture-based methods or using PCR, which has a limited screening to a few preselected known virulence genes, found specific virulence factors and antimicrobial resistance genes associated to uterine disease occurrence (23–26). Previous genomic comparisons of intrauterine *E. coli* from metritic cows from whole-genome sequencing (WGS) used low-depth sequencing from very few isolates that do not represent the genomic diversity of the organism (27). The high genetic diversity and genomic plasticity of *E. coli* are well recognized (28) and must be considered with examining genomic features for causality in disease. Therefore, genome-wide association studies including genomes from control animals and different infection sources and at a sufficient genome scale are required to identify specific adaptive traits in multi-factorial diseases *in vivo*.

Extended-spectrum beta-lactamase (ESBL) confers resistance to a wide range of beta-lactam antimicrobials, including third-generation cephalosporins, such as ceftiofur, the most used antimicrobial for metritis treatment. Livestock have been recognized as reservoirs for ESBL-producing *E. coli* and therefore have been assumed to represent a source for disseminating and spreading this important AMR genetic element to human populations (29, 30). Molecular approaches to characterize ESBL genes in dairy cattle and compare the prevalence between infectious diseases that most frequently received

antimicrobial treatment (1, 31) could guide future efforts for treatment success and AMR surveillance programs.

We hypothesized that *E. coli* isolated from postpartum dairy cows would have unique distinctive genomic characteristics, including virulence genes when compared with isolates from cows without metritis. To test this hypothesis three objectives were undertaken: (i) identify the prevalence of VF and AMR genes in *E. coli* isolated from postpartum dairy cattle from 25 commercial dairy farms in California, (ii) determine the specific genes associated with clinical disease via population comparative microbial genomics, and (iii) determine the association between the antimicrobial phenotypic resistance and the antimicrobial resistance genes identified on the *E. coli* isolated from postpartum cattle.

RESULTS

An isolate set from this study of 148 *E. coli* WGS was used for population comparisons. Out of the 148 genomes, 54 were retrieved from cows categorized based on the vaginal discharge (VD) as cows with metritis (MET), 45 from cows with purulent discharge (PUS), and 49 from cows with normal discharge (CTL) (Table 1). The genomes ranged from 4,526,522 to 6,358,509 bp assembled per genome. An additional 130 genomes were retrieved from NCBI, but 8 were excluded due to poor quality, and 2 genomes were excluded due to contamination, resulting in a total of 120 additional genomes being included in the population analysis. The genomes ranged from 4,580,125 to 5,481,723 bp assembled per genome (in Tables S1 and S2).

Comparative population genomics

The pan-genome from the 148 genomes isolated in this study was examined (Fig. S1) to find that the pan-genome was open. From this initial result, additional public WGS were used to determine the population pan-genome of 268 *E. coli* (148 from our study and 120 from NCBI). We found that this pan-genome was also open and contained 36,627 orthologous genes with 2,635 core genes and 33,992 accessory genes (softcore = 440, shell = 2,419, and cloud = 31,133) that represented the 7.2% and 92.8%, respectively (Fig. 1).

The population genome comparison of the genome distance among the 268 genome assemblies revealed two distinct groups with isolates from mastitis sources, two distinct groups with isolates from diarrhea sources, and a fifth distinct group with isolates from all clinical groups (Fig. 1C). Unexpectedly, isolates from uterine sources (metritis and no metritis) did not form an independent genomic cluster. These observations suggest that there are no specific genes in *E. coli* and, therefore, a specific *E. coli* pathotype that would increase an individual strain to more effectively result in infection of the uterus resulting in metritis in cattle.

The pan-genome analysis showed that the *E. coli* genomes from the uterus of cows with or without metritis were highly diverse and different from each other, again leaving the adaptation of *E. coli* to the uterus using genomic characterization unsupported. The Pan-GWAS analysis did not uncover any significant association between these genomes and virulence factors or antimicrobial resistance genes that allow the identification of specific or more virulent *E. coli* associated with metritis. The microbial pan-GWAS analysis determined several genomic features from the pan-genome associated with the

TABLE 1 Descriptive information for 148 postpartum dairy cows from which *E. coli* uterine isolates were retrieved in California dairy farms

| Variable	Metritis	Pus	Control
n of cows	54	45	49
DIM[a]	8 (3–21)	10 (3–21)	9.5 (3–21)
Lactation	2 (1–10)	3 (1–6)	2 (1–6)
Farms sampled	23	21	21

[a]DIM: days in milk.
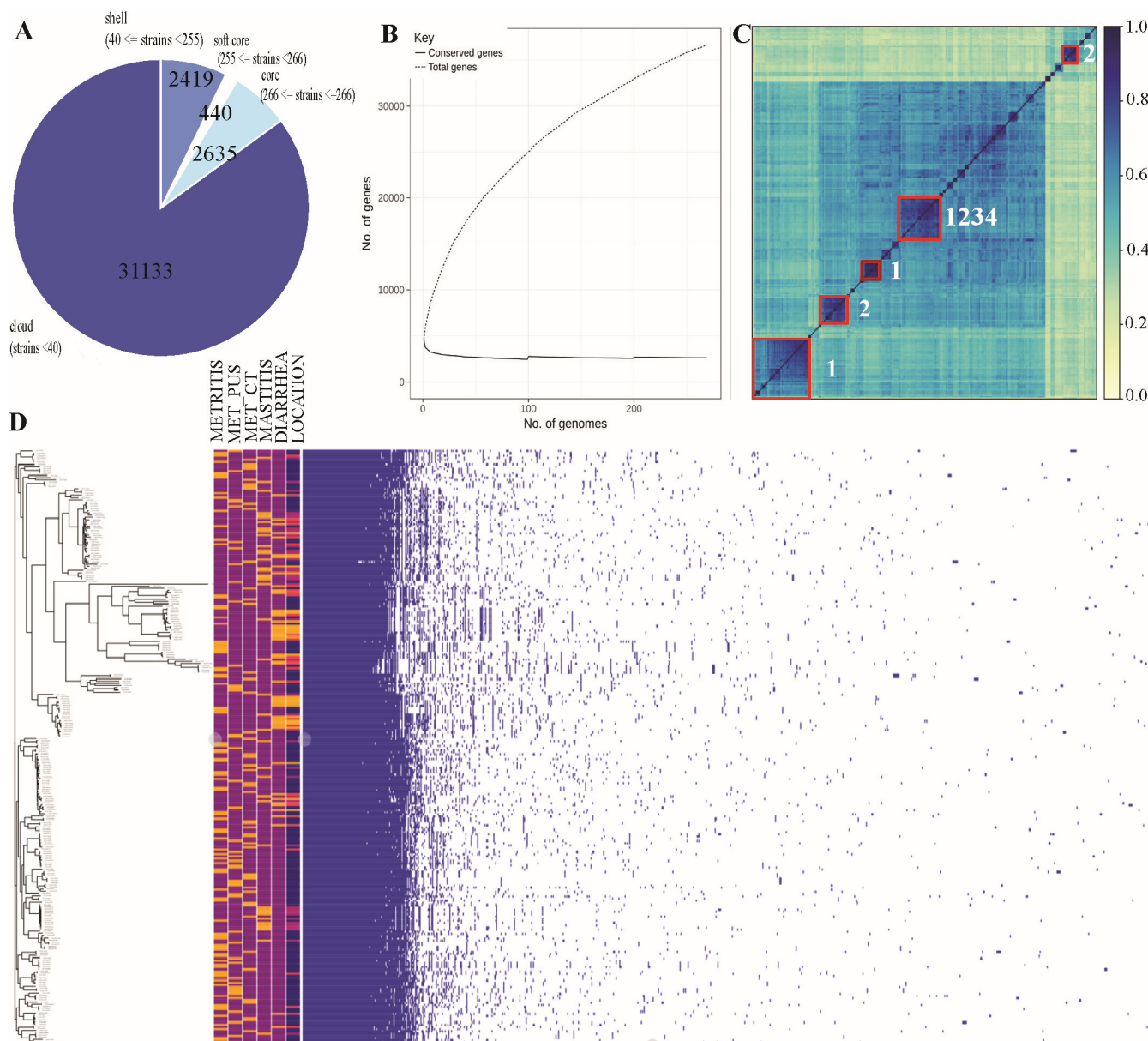
FIG 1 Pan-genome analyses of *E. coli* isolates. (A) Pie chart showing the proportion of repertoire genes in the core, soft-core, shell, and cloud of the pan-genome of the isolates (*n* = 268). (B) Pan-genome rarefaction curves showing the open pan-genome. (C) Whole-genome distance matrix depicting an all-against-all comparison of genome diversity for all isolates associated with disease phenotype. Disease phenotypes were labeled as number (1: mastitis, 2: diarrhea, 3: control metritis, and 4: metritis) (D) Gene presence-absence matrix of the gene distribution in each genome (orange: present, purple: absent), along with the metadata indicating the isolation source [metritis, metritis–purulent discharge (MET_PUS), control–metritis (MET_CT), mastitis, and diarrhea] and the geographic location (blue: United States, pink: Germany, orange: France, and yellow: Canada) of each genome.

clinical groups of diarrhea and mastitis (Table S3). Collectively, the genome diversity and pan-genome investigation did not support the theory that metritic *E. coli* are tissue adapted. With this assumption in mind, our findings do not exclude that generic *E. coli* can cause metritis in dairy cattle but instead that there is not a specific genomic marker associated with *E. coli* that was identified in cases of metritis in cattle. The large genomic diversity suggests that the disease complexity and progression are not associated with a single *E. coli* gene or genotype but instead with other factors, such as the host immune defense and uterine environmental factors or bacterial metabolic adaptation and physiological changes to stress response.

Virulence factors and antimicrobial resistance gene analysis

There was a significantly higher gene abundance of the categories of iron uptake (%, P = 0.001), fimbrial adhesins (P < 0.001), and Type II secretion system (P = 0.001) in the genomes from diarrhea compared with the other clinical groups and a significantly lower gene abundance of Type III secretion system and non-fimbrial adhesins (P = 0.001). The genomes from mastitis had a significantly higher gene abundance of genes encoding for fimbrial (P = 0.01) and non-fimbrial adhesins (P = 0.01) compared with genomes from cows with or without metritis (Fig. 2). A gene previously believed to be associated with metritis (*fimH*) (25, 26) was identified in every genome in our study as well as in all genomes from the public domain, independently of the isolation source and, therefore,
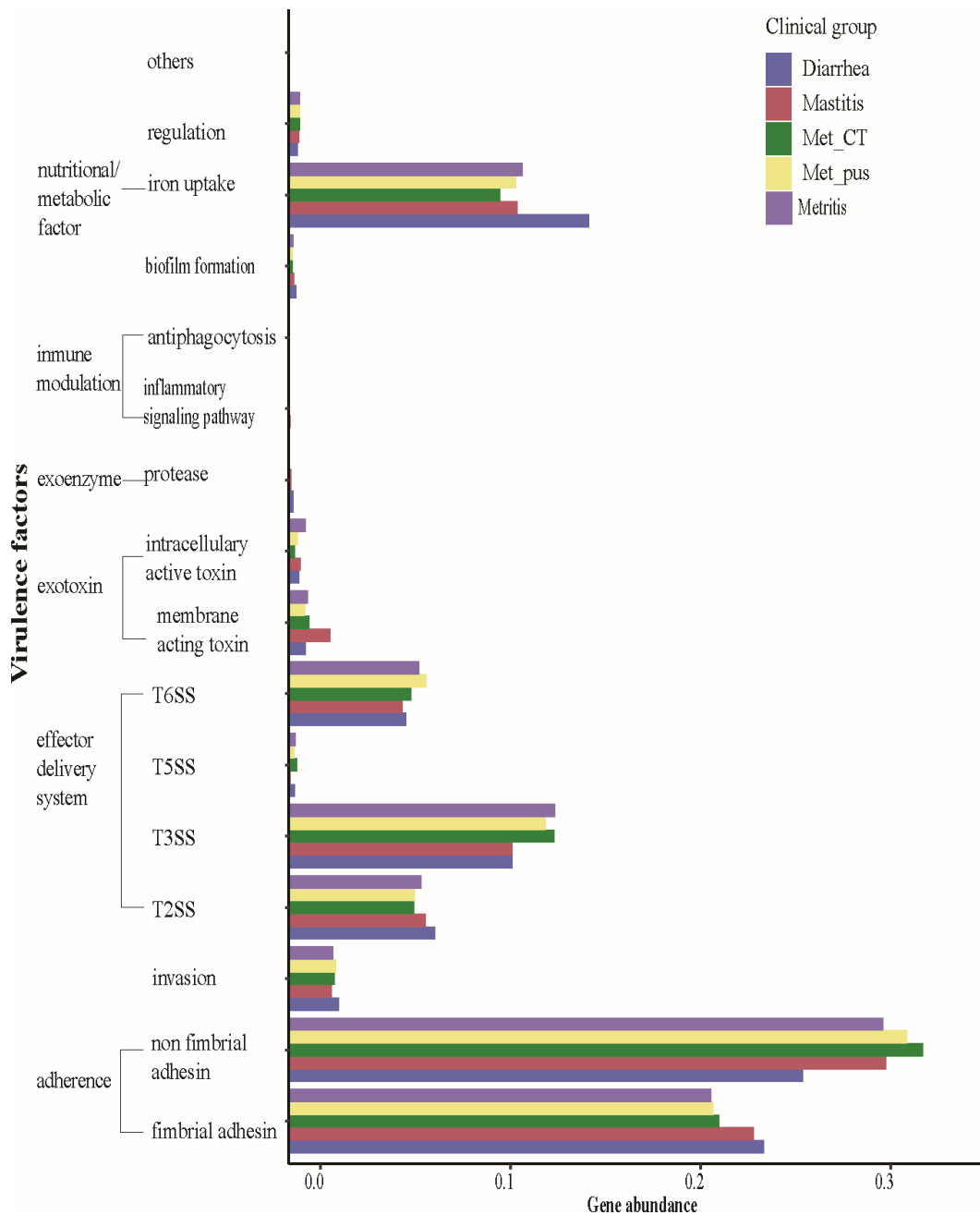
FIG 2 Virulence factor gene abundance as per health group and virulence factor category of 268 *E. coli* genomes from dairy cows classified in five different clinical groups.

not associated with the clinical group (Tables S4 and S5). Three different *fimH* alleles were detected, without any statistical association among the allelic variation of the gene and any of the health groups. There was a significantly higher gene abundance for the drug classes of aminoglycosides ($P < 0.001$), sulfonamides ($P = 0.001$), and tetracyclines ($P = 0.001$) on the genomes from diarrhea when compared with the other clinical groups (Fig. 3A). The complete results from the permutational multivariate ANOVA (PERMANOVA) analysis are shown in Table S4.

Twenty unique ESBL genes and gene variants were identified in the genomes. The prevalence of genomes carrying ESBL genes was 24.2% ($n = 60$), with a median of three (range: 1–5) unique genes detected in a single genome (Fig. 3B). The genomes with a higher abundance of ESBL genes belonged to diarrhea and metritis clinical groups. At the univariate analysis, diarrhea genomes had higher odds of carrying *TEM-105* (OR: 6, CI: 3–14, $P < 0.001$), *TEM-1B_1* (OR: 15.1, CI: 5.7–54, $P < 0.001$), and *TEM-1* (OR: 3.5, CI: 1.9–6.8, $P < 0.001$), while metritis genomes had a higher odds of carrying *CTX-M-124* (OR: 18.9, CI: 2.3–623, $P < 0.001$) when compared with the other clinical groups. When comparing the 148 isolates from this study, no significant difference in the prevalence of antimicrobial resistance or virulence factor genes between the clinical groups and drug classes or virulence factor categories was observed. The genotypic profile of the 148 *E. coli* genomes from postpartum cows showed the presence of 504 unique virulence genes and 117 antimicrobial resistance genes as per drug class and resistance mechanism (Tables S5 and S6). Additionally, 116 unique serotypes were identified (Table S7). The most abundant resistance genes found at the drug class level were aminocoumarin (23.3%), followed by beta-lactams (12.3%) and fluoroquinolones (10.6%). A total of 33.8% of isolates were multidrug resistant. The most abundant virulence genes belong to the adherence class (55%), followed by the effector delivery system class (27%) (Fig. 4).

Genotype–henotype agreement

The genotype-phenotype agreement found a variable accuracy in predicting the phenotype of the isolates having the genetic information (Table 2). Accuracy varied from 0.7% to 100%, with the lowest value being for enrofloxacin and gentamicin and the highest accuracy value for ampicillin. For most of the drugs, the high prevalence of false positives led to a low specificity for detecting genes corresponding to drugs for which they were phenotypically susceptible based on standard antimicrobial susceptibility testing methods (32).

DISCUSSION

This study describes the first population-scale ($n = 148$) genomic comparison of *E. coli* genomes collected from postpartum dairy cows from multiple farms and generated data
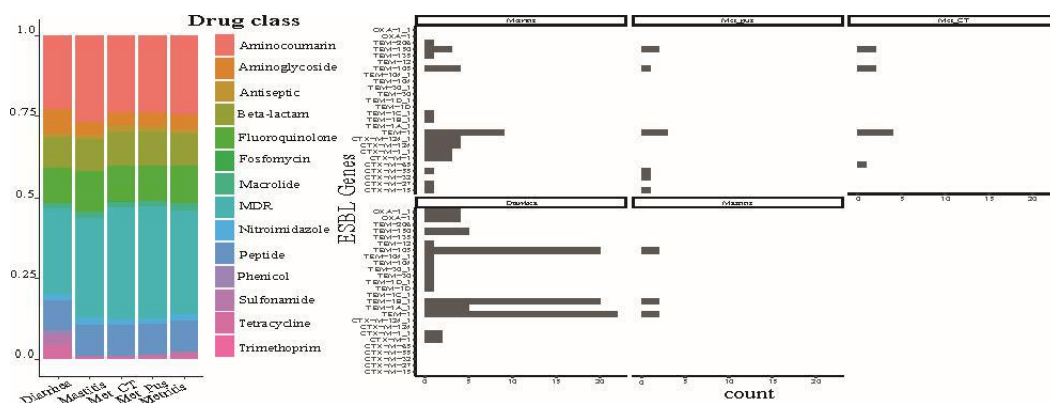
FIG 3   (A) Antimicrobial resistance gene abundance as per drug class between health groups of 268 *E. coli* genomes from dairy cows. (B) Prevalence of ESBL genes between health groups ($n = 60$).
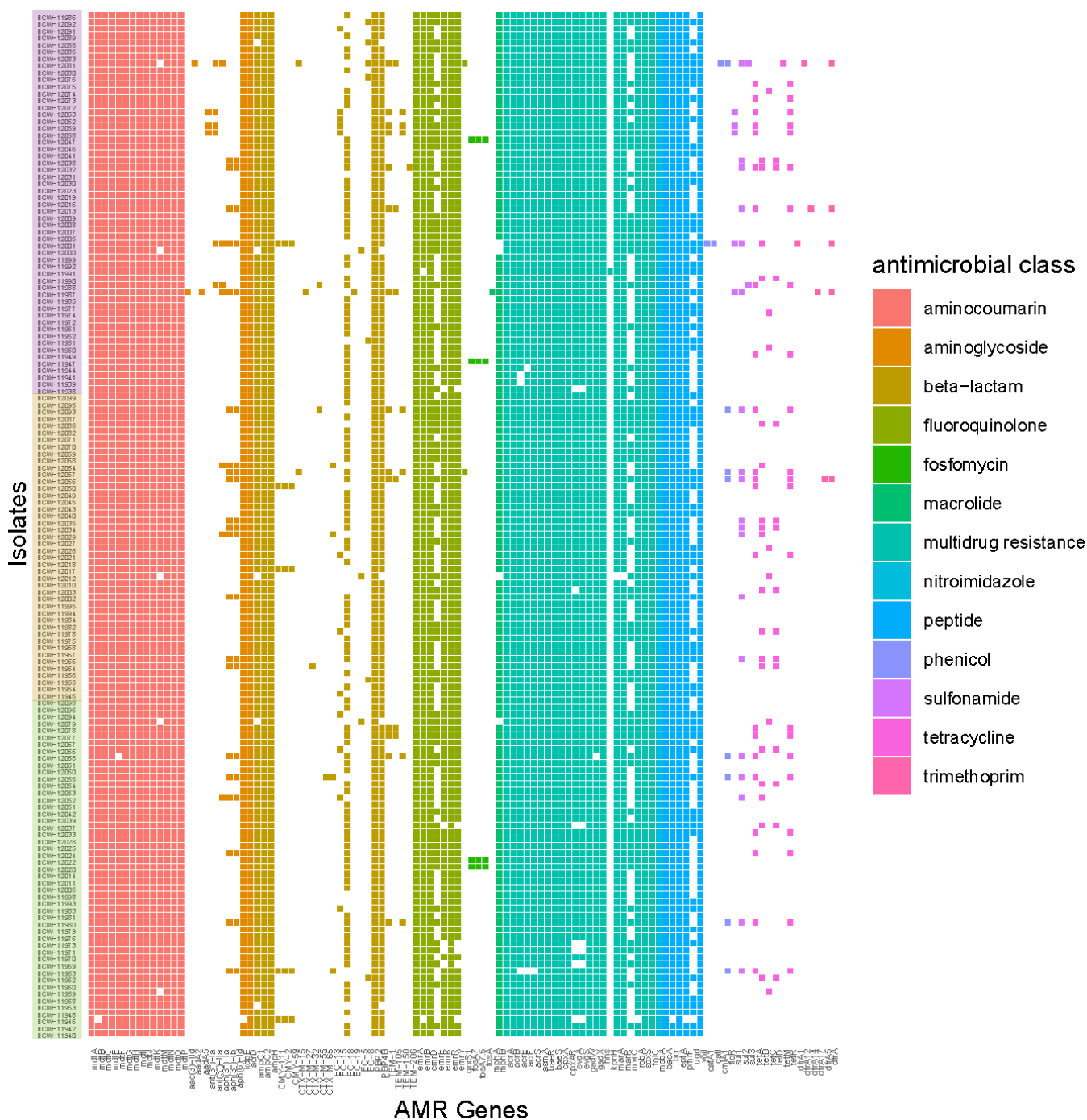
FIG 4 Presence or absence matrix of antimicrobial resistance genes by drug class and their corresponding health group (control: green, pus: orange, and metritis: purple) for 148 *E. coli* isolates from uterus of postpartum dairy cows.

that do not support previously recognized concepts of tissue adaptation and specific gene virulence factors in *E. coli* isolated from cows with metritis having unique virulence factors when compared with cows without metritis. This observation aligns with that of other organisms that have complex disease progress such as *Helicobacter pylori* where the genotype evolves during the disease (33). Furthermore, sequencing population WGS revealed that the pan-genome of uterine *E. coli* did not differentiate among the clinical groups and that there was no evidence of a specific genotype associated with metritis or uterine isolation source; therefore, the traditional EnPEC (34) pathotype was not evident

TABLE 2 Comparison of phenotypic antimicrobial susceptibility testing and genome-derived resistance prediction *for E. coli* isolates (*n* = 148)

Antibiotic	AMR genes	Susceptible phenotype		Resistant phenotype		Accuracy	Sensitivity	Specificity
		Resistant genotype FP	Susceptible genotype TN	Resistant genotype TP	Susceptible genotype FN	(TP + TN)/ total	TP/(TP + FN)	TN/(TN + FP)
Ampicillin	*ampC1* (*n* = 143), *ampC2* (*n* = 148), *ampH* (*n* = 148), CMY-111 (*n* = 5), CMY-2 (*n* = 5), CMY-59(*n* = 5), PBP2 (*n* = 148), PBP4B (*n* = 143), CTX-M-15 (*n* = 2), CTX-M-27 (*n* = 1), TEM-1 (*n* = 14), TEM-105 (*n* = 6), TEM-150 (*n* = 7), TEM-206 (*n* = 1)	0	0	148	0	1.00	1.00	0
Ceftiofur	*ampC1* (*n* = 143), *ampC2* (*n* = 148), *ampH* (*n* = 148), CTX-M-15 (*n* = 2), CTX-M-27 (*n* = 1), CTX-M-32 (*n* = 1), CTX-M-55 (*n* = 2), CTX-M-65 (*n* = 1), EC-13 (*n* = 4), EC-15 (*n* = 12), EC-18 (*n* = 106), EC-19 (*n* = 1), EC-5 (*n* = 3), EC-8 (*n* = 7), TEM-1 (*n* = 14), TEM-105 (*n* = 6), TEM-150 (*n* = 7), TEM-206 (*n* = 1)	130	0	18	0	0.12	1.00	0
Gentamicin	*aac(3)-lid* (*n* = 1), *aac(6′)-Iaa* (*n* = 1), *aac(6′)-Iy* (*n* = 1), *aadA2* (*n* = 1), *aadA5* (*n* = 1), *ant(3″)-Ia* (*n* = 3), *ant(3″)-IIa* (*n* = 7), *aph(3″)-Ia* (*n* = 3), *aph(3′)-Ib* (*n* = 21), *kdpE* (*n* = 147), *acrD* (*n* = 148)	147	0	1	0	0.007	1.00	0
Florfenicol	*catA1* (*n* = 1), *catl* (*n* = 1), *cmlA1* (*n* = 1), *floR* (*n* = 8)	2	132	7	7	0.93	0.50	0.98
Chlortetracycline	*tetA* (*n* = 22), *tetB* (*n* = 18), *tetC* (*n* = 7), *tetD* (*n* = 14),	7	97	40	4	0.92	0.90	0.93
Oxytetracycline	*tetM* (*n* = 1), *tetR* (*n* = 22)	3	93	44	8	0.92	0.84	0.96
Enrofloxacin	*emrA* (*n* = 148), *emrB* (*n* = 146), *emrD* (*n* = 148), *emrE*	147	0	1	0	0.007	1.00	0
Danofloxacin	(*n* = 65), *emrK* (*n* = 143), *emrR* (*n* = 148), *emrY* (*n* = 142), *qnrS1* (*n* = 2)	141	0	7	0	0.04	1.00	0

in the isolates from this study and it did not emerge with additional public WGS to increase the analytical power of the analysis. An explanation for these findings is that the well-recognized plasticity and genetic diversity of *E. coli* contribute to its ability to adapt to different ecological niches (28, 35, 36). A broader comparison of our metritis genomes to those from other metritis studies and isolated from animals with diarrhea and mastitis further supported the hypothesis of a non-unique genotype in the metritis isolates, revealing instead that multiple genotypes comprise the pathotype previously referred to as EnPEC (Fig. 2). Unlike other ruminant pathogens, such as *Campylobacter jejuni* which causes abortion in goats and sheep, where a single gene allele of *porA* was found to be selected over 35 years and six geographic locations to induce disease (37), we did not find any evidence of the selection or adaptation to the uterus. Our findings highlight the discriminatory power of WGS, which has been similarly used under other disease conditions to outline incongruencies when compared with more typical serotyping classification methods on genomes from avian pathogenic *E. coli*, as an example (38).

Our study found significant differences in the prevalence of virulence genes in *E. coli* from the mastitis and diarrhea groups, but not on the genomes of uterine origin. Previously conducted studies have evaluated *E. coli* virulence factors from uterine samples on postpartum dairy cows yielding contradictory results regarding their association with metritis. Most of these studies have used PCR as the detection method limiting the results to the preselected genes to be screened. Silva et al. (39) characterized the genomic profiles of *E. coli* recovered from postpartum dairy cows; in their study, 72 isolates from healthy (*n* = 35) and metritic (*n* = 37) cows within the first 42 days in milk (DIM) were screened for virulence genes using PCR, and from the 15 screened virulence genes, only 6 were identified in the samples and none of them were associated with uterine health status (39). Similarly, Sheldon et al. (34) did not find an association between uterine disease and 17 different virulence genes screened in 114 *E. coli* isolates recovered from cows that were up to 4 weeks postpartum; all isolates evaluated in

their study did not possess any of the 16 examined virulence genes, and only a gene associated with iron uptake (*fyuA*) was identified in isolates from cows with uterine infection while the gene was not identified in any of the isolates retrieved from healthy cows (34). In contrast to these studies, Bicalho et al. (26) evaluated virulence factors from *E. coli* isolated from 125 dairy cows within 3 to 7 DIM using PCR; for their study, *E. coli* isolates were screened for 32 different VF, and a significant association between six VF (*fimH*, *astA*, *cdt*, *kpsMIII*, *ibeA*, and *hlyA*) in *E. coli* and metritis was observed when compared with culture-negative cows (26). A limitation of this study arises from the study design and data analysis, where the risk of specific VFs (e.g., *Fim*H) for causing metritis was compared with that of cows with a negative intrauterine *E. coli* culture swab result (cows with no *E. coli* isolated from the sample), instead of a culture-positive sample (cows with *E. coli* without the specific VF isolated). Under these circumstances, the presence or not of *E. coli* is the only factor that can be compared when evaluating the risk for metritis. To evaluate the effect of specific VF, an appropriate reference group with an *E. coli* culture positive and VF negative is needed to compare an *E. coli* culture-positive VF present and the risk of metritis when comparing both groups, as has been done in our study.

A similar study conducted in 2012 by the same research group (25) extracted DNA directly from intrauterine swabs to evaluate the association between *E. coli* carrying only the virulence gene *fimH* and metritis using PCR. They found *E. coli* isolates carrying *fimH* during all sampling points through the study (1–3, 8–10, and 36–36 DIM) with a higher percentage during day 8, when compared with cows with PCR-negative test results for *E. coli* isolates carrying the *fimH* gene. The study also found an increased odds ratio (OR 4.7, $P < 0.01$) for metritis in cows when *E. coli* carrying *fimH* was isolated within the 1–3 DIM. A limitation of the Bicalho et al. study (25) was similar to the one outlined above for their 2010 study, where a lack of an appropriate control group (*E. coli* positive by PCR but not carrying *fim*H or other VF of interest) limits any conclusions for the association of *fim*H as the causative risk factor for metritis, and the result probably should be restricted to present or not of *E. coli* with positive test results as determined by PCR, with not conclusion alluding an association of specific VF with metritis.

A recent study by Kassé et al. (24) evaluated 371 cows within 3–20 DIM, for 40 different virulence genes using PCR (24), and identified 32 virulence genes among the isolates, with two genes being associated with metritis (*hra1* and *kpsMTII*). In contrast with the study by Bicalho et al. (26), the Kassé et al. (24) study found that *fimH* was the most abundant gene found (89%) in both metritis and healthy animals, and therefore, it was not considered a factor associated with uterine disease (24). Like the Kassé study (24), our study found *fimH* was not associated with metritis; likewise, our results go further evaluating 504 VF genes, without any specific VF associated with metritis.

FimH encodes an adhesive subunit protein of type 1 fimbriae responsible for D-mannose-sensitive adhesion, an important determinant for adhesion and subsequent colonization of mucosal surfaces (40–42). More than 85% of *E. coli* (commensal and pathogenic) expresses type 1 fimbriae (43–45). However, FimH has been recognized as an important VF for invasion of the urinary epithelium in urinary tract infections, as well as of intestinal epithelial in colitis (46) and IBD, and associated with adherence, invasion, and inflammatory response in Crohn's disease (47). Together, our findings indicate that *fim*H is a core gene that will be found in all *E. coli* isolates associated with cattle, rather than a VF that is causal to the development of metritis in dairy cows. These observations highlight the need to use appropriate control groups and population genomics to understand gene distribution in highly mutable pathogens.

Multiple factors could have influenced the difference in results between the three studies outlined above using PCR for gene detection, including study design and inclusion of control groups as previously mentioned. A more in-depth evaluation of the inconsistency of detection for important virulence genes of bacteria such as Shiga toxin-producing *E. coli* has been addressed before (48–50). Furthermore, WGS evaluation provides a much more robust method with greater comprehensive and discriminatory

power when compared with primer-dependent methods such as PCR, which has various limitations that can be overcome by new sequencing technologies (51). Allele variations in a gene can result in a gene not being detected using PCR approaches, whereas WGS would not have this limitation. Because of these advantages, WGS is considered a more effective and reliable testing approach when compared with PCR approaches (51, 52). Our study represented one of the first studies using a deep-sequencing approach from a broad source of intrauterine *E. coli* to evaluate factors associated with postpartum metritis in dairy cattle (27, 53).

Regarding antimicrobial resistance genes, our study found that the most abundant genes were found to confer resistance to aminoglycosides, tetracycline, and beta-lactams. Similarly, Jeamsripong et al. (54) also found the highest prevalence of resistance in fecal *E. coli* from dairy cattle was to tetracycline (27.5%), sulphonamide (22.5%), and aminoglycoside (20.0%). A multicentric study in Europe evaluated the genomic profile of 150 fecal *E. coli* isolates from veal calves and found that the most common AMR by drug class were against tetracycline (67%), aminoglycoside (61%), sulphonamide (58%), and beta-lactam (49%). Likewise, the most prevalent genes associated with resistance were *tetA* (56%), *aph(6)-Id* (43%), *sul2* (40%), *aadA1* (39%), and *TEM-1B* (30%) (55). Broad similarities found between our metritis genomes when compared with fecal samples reflect the findings also observed when comparing the abundance of antimicrobial resistance genes between intrauterine *E. coli* from our study and those from mastitis and diarrhea cases.

To our knowledge, only one previous study aimed to characterize the AMR genotype of *E. coli* from postpartum dairy cattle (27). This study found 71 different AMR genes in 13 *CTX-M*-positive isolates and 1 non-ESBL isolate from 4 cows with metritis. However, a small sample size, a low threshold for gene identification (>70% similarity), and a lack of a control group comparing *E. coli* from cows without metritis limit the generalization of their findings. Our study overcame this limitation by including *E. coli* from control animals and not utilizing a selective media only to characterize ESBL-positive strains, therefore representing a broader outlook of intrauterine *E. coli* from dairy cows.

The emergence and dissemination of ESBL genes have become a major concern in public health (29, 30). Although there is a knowledge gap about the directionality in the dissemination of AMR bacteria and their resistant determinants between food animals and humans (56), ESBL represents a challenge for the successful treatment of infectious diseases (57, 58). Our results showed that *TEM-1*, *TEM-105*, and *CTX-M-124* were the most abundant ESBL genes and were associated with diarrhea (*TEM-1* and *105*) or metritis (*CTX-M-124*). *TEM-1* is one of the most widely studied ESBL genes and has been found in *E. coli* distributed globally (59). TEM enzymes represent an important concern to veterinary medicine, for their effect on penicillins and early-generation cephalosporins, which are commonly used in lactating cattle (31). *CTX-M-124*, a *CTX-M-2* variant (60), was first reported in wild birds (61) and has also been described in corvids (62) and dogs and cats admitted to a veterinary teaching hospital (63). *CTX-M* enzyme family confers resistance to cefotaxime (64), a third-generation cephalosporin closely related to the primary antimicrobial use for metritis treatment (ceftiofur) (1, 65). Further research is needed to understand the role of this resistance gene and the implication of treatment success.

Our findings demonstrated a moderate overall accuracy for whole-genome sequencing to predict phenotypic resistance, which considerably varied depending on the specific antimicrobial tested. Previous studies have evaluated the correlation between WGS AMR genes and antimicrobial susceptibility, and the accuracy of predicting antimicrobial susceptibility of isolates has also reported a highly variable correlation dependent on the antimicrobial drug (66–69). A study evaluating the AMR profile of 150 fecal *E. coli* isolates from beef cattle reported a positive genotype and phenotype correlation ($r \geq 0.85$; $P < 0.05$) of the isolates. The study also found that most of the discordant results came from isolates harboring genes responsible for aminoglycoside and sulfonamide resistance (70). Similarly, reference 54 described the

antimicrobial resistance profiles of 40 fecal *E. coli* from dairy cattle at different production stages on a commercial farm. Results from this study found a high agreement (85.71%–97.50%) between the resistance phenotype and the presence of resistance genes. However, kappa statistics varied greatly between antimicrobials, showing a weak level of agreement between aminoglycoside resistance genes and streptomycin (39.5%) and trimethoprim and trimethoprim–sulfamethoxazole (48.9%) resistance and a moderate agreement for beta-lactamases ampicillin (78.7%) and tetracyclines (60.8%). Genotype–phenotype correlation variation among studies might be influenced by several factors such as small sample size, databases included for AMR gene identification, thresholds used for gene identity and coverage, and MIC values used for susceptibility classification of isolates. Results of genotypes in the current work and phenotypes in our previous work (32) on the same bacterial isolates allowed us to better understand the resistance of uterine *E. coli* on dairy farms.

Different factors involving the host genetics and immune response, the epithelial barrier disruption at calving, and the uterine microbiome warrant further research to better understand the pathogenesis of the disease. Changes in microbiota after calving and the reduced diversity have been described before in metric cows compared with healthy cows (11, 16, 17). Microbiome studies have found that the dynamics of uterine microbiota differs between cows with metritis and control cows and have attributed the disease as a consequence of uterine dysbiosis (71–73). This dysbiosis favored the overgrowth of facultative anaerobic bacteria such as *Fusobacterium* spp. and *E. coli*, which could explain the higher prevalence of *E. coli* isolated from cows with metritis than from control cows or cows with purulent vaginal discharge found in the present study.

E. coli has been identified as the causal agent of metritis using culture-based methods, and the disease has been successfully replicated in *in vivo* models using *E. coli* or in combination with other bacteria such as *T. pyogenes* (19, 20, 74). Microbiome studies have found a variable association between *E. coli* and metritis when compared with control cows (72, 73). Differences in disease development among these groups may benefit from a better understanding of the evolution of pathogenic *E. coli*. The high genetic plasticity along with the horizontal gene transfer of pathogenicity islands favor the rapid adaptation of *E. coli* to new ecological niches (75, 76), allowing commensal strains to have pathogenic potential or act as genetic reservoirs for virulence factors and AMR genes (75). Finally, understanding the host factors relevant to the development of the disease such as the metabolic status of the cows at calving and the negative effect on their immune system, as well as the interaction of the host immune response with pathogenic bacteria, may help understand the differences in prevalence and disease severity among animals. The role of *E. coli* has been discussed in human diseases in which pathology is not completely understood yet either, like in ulcerative colitis, and the bacteria's ability to cause disease depends on the host immune status, the intestinal microbiota, and the intestinal immune response (77). A better understanding of the interaction of *E. coli* in the uterine immune response in postpartum cows may help elucidate the differences in disease development when the bacteria is present.

Conclusion

An *E. coli*-specific genotype was not found to be associated with metritis in this study. This undermines the concept that *E. coli* are tissue adapted as was found with other dairy cattle *E. coli*-related diseases. In contrast, a high genetic diversity among the isolates from uterine sources (metritis, PUS, and control) is in conflict with the current classification for uterine-specific *E. coli* niche adaptation theory. A virulence factor previously associated with *E. coli* in metritis (*fim*H) was equally prevalent between all isolation sources and found to be a core gene in the population pan-genome comparison and was not specifically associated with the development of metritis. Although three different *fimH* alleles were identified across the *E. coli* genomes, there was not any association between allele variation and metritis. ESBL genes were more abundant in genomes from diarrhea and metritis clinical groups. Genomes from the metritis clinical

group had higher odds of carrying the *CTX-M-124* ESBL gene, while genomes from the diarrhea group had higher odds of carrying *TEM*-105, *TEM-1B_1*, and *TEM-1* ESBL genes. Our findings generated novel data that provide significant advancement to the field, for redefining the traditional definition of uterine-specific *E. coli* and highlighting the need for continuous revision of previous concepts of metritis pathogenesis using new deep-sequencing technologies. Further research that investigates the uterine micro-biome in cows with and without metritis, as well as the role of the host immune response in disease development and postpartum involution, may help better understand metritis pathophysiology and tailor interventions for disease diagnosis and successful treatment.

MATERIALS AND METHODS

Sampling and study design

All procedures were approved by The University of California Institutional Animal Care and Use Committee (no. 20620). The study was conducted between September 2018 and November 2019.

Bacterial isolates (*n* = 163) were collected as part of a larger cross-sectional study that collected uterine swabs from postpartum cows between 3 and 21 days in milk using a convenience sample of 25 commercial dairy farms from the Sacramento and San Joaquin Valleys in California (32). Three clinical groups were defined based on VD characteristics as metritis discharge: watery, reddish, or brownish, and fetid, purulent discharge: non-fetid purulent or mucopurulent vaginal discharge, and normal discharge: clear lochia, clear mucus, or no vaginal discharge (Fig. 5).

DNA extraction, library preparation, and whole-genome sequencing

Frozen bacterial stocks were used to inoculate a CHROMagar- *E. coli* selective plate (CHROMagar Microbiology, Paris, France) and were incubated aerobically at 37°C for 24 hours. A single, well-isolated colony was subcultured in 1.5 mL of sterile Luria–Bertani broth (Difco; Becton, Dickinson, and Company, Sparks, MD, USA) at 37°C for 24 hours. Culture tubes were centrifuged at 10,000 × *g* for 3 min at room temperature (15°C–25°C), and the pellet was then resuspended in 180 µL Buffer ATL.

DNA was extracted using DNeasy Blood & Tissue (QIAGEN, Hilden, Germany), following the manufacturer's instructions, and DNA concentration was measured at 260 nm using a NanoDrop One^c Microvolume UV-VIS spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). Additionally, for samples with a concentration < 50 ng/µL or A260/260 < 2.0, DNA was purified using Genomic DNA Clean & Concentra-tor-25 (Zymo Research Corp., Irvine, CA, USA).

DNA library preparation was conducted in the laboratory of Dr. Bart Weimer (UC Davis). DNA was analyzed on the Agilent 2200 TapeStation System using the Genomic DNA ScreenTape assay for the integrity of gDNA. Libraries were constructed using the KAPA HyperPlus Library Preparation Kit (Roche, Indianapolis, IN, USA), as previously described (78, 79). Whole genome sequencing was done using the Illumina HiSeq X platform with PE150 (Illumina Inc., San Diego, CA, USA) (80).

Sequence assembly, annotation, and pan-genome analyses

The genome sequences were assessed for sequencing depth (>20× estimate), checked for quality using FastQC (v0.11.9) (81), and trimmed using Trimmomatic (v0.39) (82). Sequences were assembled using Shovil (v1.0.4) (83), checked for quality, size (4.5–6.5Mbp genome), completeness (>95% estimate), and contamination (<10% estimate) using CheckM (84), and assessed for approximate genera and species and further identity test for possible contamination using Kraken (85–90). Sixteen sequences that did not meet quality criteria were removed from downstream analysis.

The comparative WGS analyses were extended to include 130 public genomes of *Escherichia coli* of clinical importance from cows with diarrhea (*n* = 50, BioProject:
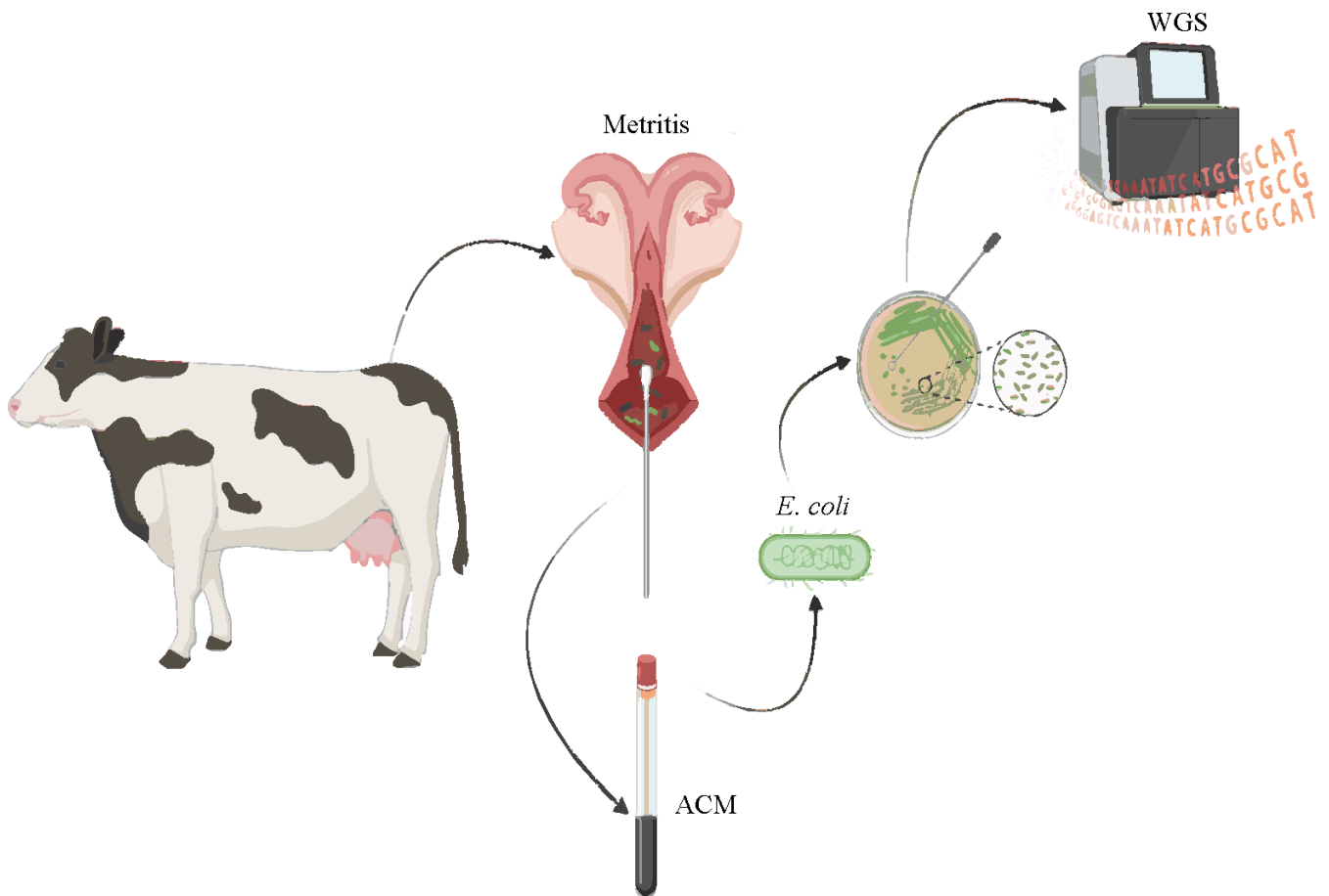
FIG 5 Graphical summary depicting the study design and methodology used to isolate and sequence intrauterine *E. coli* (*n* = 163) from cows with and without metritis.

PRJEB32666), cows with clinical mastitis (*n* = 50, BioProject: PRJNA612640), and cows with metritis (*n* = 30, BioProject: PRJNA298331); these genomes were combined with the 148 *E. coli* from our study to examine their genome relatedness, following the same procedures described above. Genome distance metrics were determined using all-by-all whole genome comparisons using Sourmash (v3.2.3). The genome distance metric was calculated using genome-wide k-mer signatures, using a k-mer size of 31 with sketch sizes scaled to 100,000/megabase. Pairwise comparisons were visualized as an all-against-all heatmap (91).

Core and accessory genes were annotated using Prokka (version v.1.14.6) (92). Pan-genome comparisons, identifying gene clusters and the core genes, were conducted using Roary (3.12.0) using 95% amino acid sequence identity (93) and visualized using Phandango (94) with the associated metadata. Gene associations, metadata, and phenotypes were associated using Scoary 1.6.12 (95). Finally, variant calling was done with the reference sequence *E. coli* ECC-1470 with Snippy (version v. 4.6.0) with default settings (96).

Genomic assessment of virulence factors and antibiotic resistance genes

Virulence factors and antimicrobial resistance genes were analyzed in every genome using ABRicate (version 1.0.0) (97). Virulence factor genes were screened against the virulence factor database (VFDB), and ecoli_VF. Antibiotic resistance genes were screened against the Comprehensive Antibiotic Resistance Database (CARD), the Antibiotic Resistance Gene-Annotation (ARG-ANNOT), MEGARes, ResFinder, and the

National Database of Antibiotic-Resistant Organisms (NDARO). The AMR genes were classified according to resistance mechanisms and drug class using the CARD database and manual curation (98). The VF genes were classified according to the classification scheme proposed by the VFDB (99). Both AMR and VF determinants were retained for analysis if they fit the minimum criteria of 90% identity and coverage. Serotyping was done using the ECOH database (100).

Genotype–phenotype correlation for antimicrobial resistance

Genotype–phenotype correlation (GPC) analysis was performed using broth microdilution antimicrobial susceptibility test results (32) in combination with the WGS-informed AMR analysis. These assays were compared with the occurrence of known AMR genes with resistance to the respective drug. False-positive, false-negative, sensitivity, specificity, and accuracy of the GPC were calculated as previously described (66).

Statistical analysis

Data analysis was conducted using RStudio (version 4.1.2). Microbial pan-GWAS analysis to identify significant gene associations was conducted using Scoary, and P values were adjusted using Bonferroni correction (95). Descriptive statistics were used to examine the distribution of AMR and VF genes in *E. coli* between the clinical groups. To investigate differences between the drug classes or the virulence factor category among the clinical groups (control, pus, metritis, diarrhea, and mastitis), PERMANOVA was conducted using the vegan package (101). Post hoc pairwise comparisons with Bonferroni correction were calculated for the categories with significant differences. Univariate analyses were conducted to evaluate the association between the prevalence of ESBL genes and the clinical group. Finally, to assess the agreement between *E. coli* isolates having a specific genotype and the corresponding resistant phenotype, the accuracy, sensitivity, and specificity were calculated. Statistical significance was set for all tests at $P \leq 0.05$.

AUTHOR AFFILIATIONS

[1]Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, California, USA
[2]Department of Population Health and Reproduction, 100K Pathogen Genome Project, School of Veterinary Medicine, University of California, Davis, California, USA
[3]Veterinary Medicine Teaching and Research Center, School of Veterinary Medicine, University of California, Tulare, California, USA
[4]Cooperative Extension, Division of Agriculture and Natural Resources, University of California, Orland, California, USA

AUTHOR ORCIDs

Fabio S. Lima http://orcid.org/0000-0001-8377-6469
Bart C. Weimer http://orcid.org/0000-0002-7471-1978
Richard V. Pereira http://orcid.org/0000-0003-2028-8761

DATA AVAILABILITY

All raw genome sequences generated in this study are available at the 100K Pathogen Genome Project BioProject (NCBI PRJNA186441) under BioProject accession number PRJNA1095494.

ETHICS APPROVAL

All procedures were approved by The University of California Institutional Animal Care and Use Committee (no. 20620).

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Supplemental Tables (mBio01027-24-s0001.xlsx). WGS tables.

REFERENCES

1. USDA. 2014. Dairy 2014: health and management practices on U.S. dairy operations. # 696.0218; 2018. Fort Collins, CO USDA–APHIS–VS–CEAH–NAHMS
2. Sheldon IM, Lewis GS, LeBlanc S, Gilbert RO. 2006. Defining postpartum uterine disease in cattle. Theriogenology 65:1516–1530. https://doi.org/10.1016/j.theriogenology.2005.08.021
3. Pérez-Báez J, Silva TV, Risco CA, Chebel RC, Cunha F, De Vries A, Santos JEP, Lima FS, Pinedo P, Schuenemann GM, Bicalho RC, Gilbert RO, Rodrigez-Zas S, Seabury CM, Rosa G, Thatcher WW, Galvão KN. 2021. The economic cost of metritis in dairy herds. J Dairy Sci 104:3158–3168. https://doi.org/10.3168/jds.2020-19125
4. Lima FS, Vieira-Neto A, Snodgrass JA, De Vries A, Santos JEP. 2019. Economic comparison of systemic antimicrobial therapies for metritis in dairy cows. J Dairy Sci 102:7345–7358. https://doi.org/10.3168/jds.2018-15383
5. Haimerl P, Heuwieser W. 2014. Invited review: antibiotic treatment of metritis in dairy cows: a systematic approach. J Dairy Sci 97:6649–6661. https://doi.org/10.3168/jds.2014-8462
6. Sheldon IM, Cronin JG, Bromfield JJ. 2019. Tolerance and innate immunity shape the development of postpartum uterine disease and the impact of endometritis in dairy cattle. Annu Rev Anim Biosci 7:361–384. https://doi.org/10.1146/annurev-animal-020518-115227
7. Garzon A, Habing G, Lima F, Silva-Del-Rio N, Samah F, Pereira R. 2022. Defining clinical diagnosis and treatment of puerperal metritis in dairy cows: a scoping review. J Dairy Sci 105:3440–3452. https://doi.org/10.3168/jds.2021-21203
8. Santos TMA, Bicalho RC. 2012. Diversity and succession of bacterial communities in the uterine fluid of postpartum metritic, endometritic and healthy dairy cows. PLoS ONE 7:e53048. https://doi.org/10.1371/journal.pone.0053048
9. Laguardia-Nascimento M, Branco K, Gasparini MR, Giannattasio-Ferraz S, Leite LR, Araujo FMG, Salim A de M, Nicoli JR, de Oliveira GC, Barbosa-Stancioli EF. 2015. Vaginal microbiome characterization of Nellore cattle using metagenomic analysis. PLoS ONE 10:e0143294. https://doi.org/10.1371/journal.pone.0143294
10. Wang Y, Wang J, Li H, Fu K, Pang B, Yang Y, Liu Y, Tian W, Cao R. 2018. Characterization of the cervical bacterial community in dairy cows with metritis and during different physiological phases. Theriogenology 108:306–313. https://doi.org/10.1016/j.theriogenology.2017.12.028
11. Jeon SJ, Vieira-Neto A, Gobikrushanth M, Daetz R, Mingoti RD, Parize ACB, de Freitas SL, da Costa ANL, Bicalho RC, Lima S, Jeong KC, Galvão KN. 2015. Uterine microbiota progression from calving until establishment of metritis in dairy cows. Appl Environ Microbiol 81:6324–6332. https://doi.org/10.1128/AEM.01753-15
12. Karstrup CC, Klitgaard K, Jensen TK, Agerholm JS, Pedersen HG. 2017. Presence of bacteria in the endometrium and placentomes of pregnant cows. Theriogenology 99:41–47. https://doi.org/10.1016/j.theriogenology.2017.05.013
13. Moore SG, Ericsson AC, Poock SE, Melendez P, Lucy MC. 2017. Hot topic: 16S rRNA gene sequencing reveals the microbiome of the virgin and pregnant bovine uterus. J Dairy Sci 100:4953–4960. https://doi.org/10.3168/jds.2017-12592
14. Ruder CA, Sasser RG, Williams RJ, Ely JK, Bull RC, Butler JE. 1981. Uterine infections in the postpartum cow: II. Possible synergistic effect of *Fusobacteriumnecrophorum* and *Corynebacteriumpyogenes*. Theriogenology 15:573–580. https://doi.org/10.1016/0093-691X(81)90060-1
15. Sheldon IM, Noakes DE, Rycroft AN, Dobson H. 2004. Effect of intrauterine administration of oestradiol on postpartum uterine bacterial infection in cattle. Anim Reprod Sci 81:13–23. https://doi.org/10.1016/j.anireprosci.2003.08.010
16. Galvão KN, Higgins CH, Zinicola M, Jeon SJ, Korzec H, Bicalho RC. 2019. Effect of pegbovigrastim administration on the microbiome found in the vagina of cows postpartum. J Dairy Sci 102:3439–3451. https://doi.org/10.3168/jds.2018-15783
17. Jeon SJ, Cunha F, Ma X, Martinez N, Vieira-Neto A, Daetz R, Bicalho RC, Lima S, Santos JEP, Jeong KC, Galvão KN. 2016. Uterine microbiota and immune parameters associated with fever in dairy cows with metritis. PLoS ONE 11:e0165740. https://doi.org/10.1371/journal.pone.0165740
18. Amos MR, Healey GD, Goldstone RJ, Mahan SM, Düvel A, Schuberth H-J, Sandra O, Zieger P, Dieuzy-Labaye I, Smith DGE, Sheldon IM. 2014. Differential endometrial cell sensitivity to a cholesterol-dependent cytolysin links *Trueperella pyogenes* to uterine disease in cattle. Biol Reprod 90:54. https://doi.org/10.1095/biolreprod.113.115972
19. Piersanti RL, Zimpel R, Molinari PCC, Dickson MJ, Ma Z, Jeong KC, Santos JEP, Sheldon IM, Bromfield JJ. 2019. A model of clinical endometritis in holstein heifers using pathogenic *Escherichia coli* and *Trueperella pyogenes*. J Dairy Sci 102:2686–2697. https://doi.org/10.3168/jds.2018-15595
20. Silva JCC, Siqueira LC, Rodrigues MX, Zinicola M, Wolkmer P, Pomeroy B, Bicalho RC. 2023. Intrauterine infusion of a pathogenic bacterial cocktail is associated with the development of clinical metritis in

postpartum multiparous Holstein cows. J Dairy Sci 106:607–623. https://doi.org/10.3168/jds.2022-21954

21. Sheldon IM, Rycroft AN, Dogan B, Craven M, Bromfield JJ, Chandler A, Roberts MH, Price SB, Gilbert RO, Simpson KW. 2010. Specific strains of *Escherichia coli* are pathogenic for the endometrium of cattle and cause pelvic inflammatory disease in cattle and mice. PLoS ONE 5:e9192. https://doi.org/10.1371/journal.pone.0009192

22. Williams EJ, Fischer DP, Noakes DE, England GCW, Rycroft A, Dobson H, Sheldon IM. 2007. The relationship between uterine pathogen growth density and ovarian function in the postpartum dairy cow. Theriogenology 68:549–559. https://doi.org/10.1016/j.theriogenology.2007.04.056

23. Ma Z, Ginn A, Kang M, Galvão KN, Jeong KC. 2018. Genomic and virulence characterization of intrauterine pathogenic *Escherichia coli* with multi-drug resistance isolated from cow uteri with metritis. Front Microbiol 9:3137. https://doi.org/10.3389/fmicb.2018.03137

24. Kassé FNN, Fairbrother JMM, Dubuc J. 2016. Relationship between *Escherichia coli* virulence factors and postpartum metritis in dairy cows. J Dairy Sci 99:4656–4667. https://doi.org/10.3168/jds.2015-10094

25. Bicalho MLS, Machado VS, Oikonomou G, Gilbert RO, Bicalho RC. 2012. Association between virulence factors of *Escherichia coli*, *Fusobacterium necrophorum*, and *Arcanobacterium pyogenes* and uterine diseases of dairy cows. Vet Microbiol 157:125–131. https://doi.org/10.1016/j.vetmic.2011.11.034

26. Bicalho RC, Machado VS, Bicalho MLS, Gilbert RO, Teixeira AGV, Caixeta LS, Pereira RVV. 2010. Molecular and epidemiological characterization of bovine intrauterine *Escherichia coli*. J Dairy Sci 93:5818–5830. https://doi.org/10.3168/jds.2010-3550

27. Ma Z, Ginn A, Kang M, Galvão KN, Jeong KC. 2018. Genomic and virulence characterization of Intrauterine pathogenic *Escherichia coli* with multi-drug resistance isolated from cow uteri with metritis. Front Microbiol 9:3137. https://doi.org/10.3389/fmicb.2018.03137

28. Tenaillon O, Skurnik D, Picard B, Denamur E. 2010. The population genetics of commensal *Escherichia coli*. Nat Rev Microbiol 8:207–217. https://doi.org/10.1038/nrmicro2298

29. Müller A, Stephan R, Nüesch-Inderbinen M. 2016. Distribution of virulence factors in ESBL-producing *Escherichia coli* isolated from the environment, livestock, food and humans. Sci Total Environ 541:667–672. https://doi.org/10.1016/j.scitotenv.2015.09.135

30. Dahms C, Hübner N-O, Kossow A, Mellmann A, Dittmann K, Kramer A. 2015. Occurrence of ESBL-producing *Escherichia coli* in livestock and farm workers in Mecklenburg-Western Pomerania, Germany. PLoS ONE 10:e0143326. https://doi.org/10.1371/journal.pone.0143326

31. United States Department of Agriculture (USDA), National Health Monitoring System (NAHMS). 2008. Antibiotic use on U.S. dairy operations, 2002 and 2007

32. Basbas C, Garzon A, Silva-Del-Rio N, Byrne BA, Karle B, Aly SS, Champagne JD, Williams DR, Lima FS, Machado VS, Pereira RV. 2022. Evaluation of antimicrobial resistance and risk factors for recovery of intrauterine *Escherichia coli* from cows with metritis on California commercial dairy farms. Sci Rep 12:13937. https://doi.org/10.1038/s41598-022-18347-w

33. Draper JL, Hansen LM, Bernick DL, Abedrabbo S, Underwood JG, Kong N, Huang BC, Weis AM, Weimer BC, van Vliet AHM, Pourmand N, Solnick JV, Karplus K, Ottemann KM. 2017. Fallacy of the unique genome: sequence diversity within single *Helicobacter pylori* strains. mBio 8:e02321-16. https://doi.org/10.1128/mBio.02321-16

34. Sheldon IM, Rycroft AN, Dogan B, Craven M, Bromfield JJ, Chandler A, Roberts MH, Price SB, Gilbert RO, Simpson KW. 2010. Specific strains of *Escherichia coli* are pathogenic for the endometrium of cattle and cause pelvic inflammatory disease in cattle and mice. PLoS ONE 5:e9192. https://doi.org/10.1371/journal.pone.0009192

35. Montealegre MC, Talavera Rodríguez A, Roy S, Hossain MI, Islam MA, Lanza VF, Julian TR. 2020. High genomic diversity and heterogenous origins of pathogenic and antibiotic-resistant *Escherichia coli* in household settings represent a challenge to reducing transmission in low-income settings. mSphere 5:e00704-19. https://doi.org/10.1128/mSphere.00704-19

36. Vila J, Sáez-López E, Johnson JR, Römling U, Dobrindt U, Cantón R, Giske CG, Naas T, Carattoli A, Martínez-Medina M, Bosch J, Retamar P, Rodríguez-Baño J, Baquero F, Soto SM. 2016. *Escherichia coli*: an old

friend with new tidings. FEMS Microbiol Rev 40:437–463. https://doi.org/10.1093/femsre/fuw005

37. Bandoy DDR, Weimer BC. 2020. Biological machine learning combined with *Campylobacter* population genomics reveals virulence gene allelic variants cause disease. Microorganisms 8:549. https://doi.org/10.3390/microorganisms8040549

38. Mehat JW, van Vliet AHM, La Ragione RM. 2021. The avian pathogenic *Escherichia coli* (APEC) pathotype is comprised of multiple distinct, independent genotypes. Avian Pathol 50:402–416. https://doi.org/10.1080/03079457.2021.1915960

39. Silva E, Leitão S, Tenreiro T, Pomba C, Nunes T, Lopes da Costa L, Mateus L. 2009. Genomic and phenotypic characterization of *Escherichia coli* isolates recovered from the uterus of puerperal dairy cows. J Dairy Sci 92:6000–6010. https://doi.org/10.3168/jds.2009-2358

40. Johnson JR. 1991. Virulence factors in *Escherichia coli* urinary tract infection. Clin Microbiol Rev 4:80–128. https://doi.org/10.1128/CMR.4.1.80

41. Klemm P, Schembri MA. 2000. Bacterial adhesins: function and structure. Int J Med Microbiol 290:27–35. https://doi.org/10.1016/S1438-4221(00)80102-2

42. Krogfelt KA, Bergmans H, Klemm P. 1990. Direct evidence that the FimH protein is the mannose-specific adhesin of *Escherichia coli* type 1 fimbriae. Infect Immun 58:1995–1998. https://doi.org/10.1128/iai.58.6.1995-1998.1990

43. Arthur M, Johnson CE, Rubin RH, Arbeit RD, Campanelli C, Kim C, Steinbach S, Agarwal M, Wilkinson R, Goldstein R. 1989. Molecular epidemiology of adhesin and hemolysin virulence factors among uropathogenic *Escherichia coli*. Infect Immun 57:303–313. https://doi.org/10.1128/iai.57.2.303-313.1989

44. Sokurenko EV, Chesnokova V, Dykhuizen DE, Ofek I, Wu XR, Krogfelt KA, Struve C, Schembri MA, Hasty DL. 1998. Pathogenic adaptation of *Escherichia coli* by natural variation of the FimH adhesin. Proc Natl Acad Sci U S A 95:8922–8926. https://doi.org/10.1073/pnas.95.15.8922

45. Schembri MA, Sokurenko EV, Klemm P. 2000. Functional flexibility of the FimH adhesin: insights from a random mutant library. Infect Immun 68:2638–2646. https://doi.org/10.1128/IAI.68.5.2638-2646.2000

46. Zhang W, An E-K, Kim S-J, Park H-B, Lee PCW, Jin J-O. 2023. *Escherichia coli* adhesion protein FimH exacerbates colitis via CD11b+CD103-dendritic cell activation. Front Immunol 14:1284770. https://doi.org/10.3389/fimmu.2023.1284770

47. Reinisch W, Hébuterne X, Buisson A, Schreiber S, Desreumaux P, Primas C, Paillarse J, Chevalier G, Bonny C. 2022. Safety, pharmacokinetic, and pharmacodynamic study of sibofimloc, a novel FimH blocker in patients with active Crohn's disease. J Gastroenterol Hepatol 37:832–840. https://doi.org/10.1111/jgh.15828

48. Stanford K, Reuter T, Hallewell J, Tostes R, Alexander TW, McAllister TA. 2018. Variability in characterizing *Escherichia coli* from cattle feces: a cautionary tale. Microorganisms 6:74. https://doi.org/10.3390/microorganisms6030074

49. Castro VS, Ortega Polo R, Figueiredo EE de S, Bumunange EW, McAllister T, King R, Conte-Junior CA, Stanford K. 2021. Inconsistent PCR detection of shiga toxin-producing *Escherichia coli*: insights from whole genome sequence analyses. PLoS ONE 16:e0257168. https://doi.org/10.1371/journal.pone.0257168

50. Ferdous M, Zhou K, de Boer RF, Friedrich AW, Kooistra-Smid AMD, Rossen JWA. 2015. Comprehensive characterization of *Escherichia coli* O104: H4 isolated from patients in the Netherlands. Front Microbiol 6:1348. https://doi.org/10.3389/fmicb.2015.01348

51. Goodwin S, McPherson JD, McCombie WR. 2016. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17:333–351. https://doi.org/10.1038/nrg.2016.49

52. Slatko BE, Gardner AF, Ausubel FM. 2018. Overview of next generation sequencing technologies. Curr Protoc Mol Biol 122:e59. https://doi.org/10.1002/cpmb.59

53. Jones K, Cunha F, Jeon SJ, Pérez-Báez J, Casaro S, Fan P, Liu T, Lee S, Jeong KC, Yang Y, Galvão KN. 2022. Tracing the source and route of uterine colonization by exploring the genetic relationship of *Escherichia coli* isolated from the reproductive and gastrointestinal tract of dairy cows. Vet Microbiol 266:109355. https://doi.org/10.1016/j.vetmic.2022.109355

54. Jeamsripong S, Li X, Aly SS, Su Z, Pereira RV, Atwill ER. 2021. Antibiotic resistance genes and associated phenotypes in *Escherichia coli* and *Enterococcus* from cattle at different production stages on a dairy farm in central California. Antibiotics (Basel) 10:1042. https://doi.org/10.3390/antibiotics10091042

55. Leekitcharoenphon P, Johansson MHK, Munk P, Malorny B, Skarżyńska M, Wadepohl K, Moyano G, Hesp A, Veldman KT, Bossers A, Zając M, Wasyl D, Sanders P, Gonzalez-Zorn B, Brouwer MSM, Wagenaar JA, Heederik DJJ, Mevius D, Aarestrup FM, EFFORT Consortium. 2021. Genomic evolution of antimicrobial resistance in *Escherichia coli*. Sci Rep 11:15108. https://doi.org/10.1038/s41598-021-93970-7

56. Muloi D, Ward MJ, Pedersen AB, Fèvre EM, Woolhouse MEJ, van Bunnik BAD. 2018. Are food animals responsible for transfer of antimicrobial-resistant *Escherichia coli* or their resistance determinants to human populations? A systematic review. Foodborne Pathog Dis 15:467–474. https://doi.org/10.1089/fpd.2017.2411

57. Søraas A, Sundsfjord A, Jørgensen SB, Liestøl K, Jenum PA. 2014. High rate of per oral mecillinam treatment failure in community-acquired urinary tract infections caused by ESBL-producing *Escherichia coli*. PLoS One 9:e85889. https://doi.org/10.1371/journal.pone.0085889

58. Russo A, Berruti M, Giacobbe DR, Vena A, Bassetti M. 2021. Recent molecules in the treatment of severe infections caused by ESBL-producing bacteria. Expert Rev Anti Infect Ther19:983–991. https://doi.org/10.1080/14787210.2021.1874918

59. Salverda MLM, De Visser JAGM, Barlow M. 2010. Natural evolution of TEM-1 β-lactamase: experimental reconstruction and clinical relevance. FEMS Microbiol Rev 34:1015–1036. https://doi.org/10.1111/j.1574-6976.2010.00222.x

60. Rodríguez MM, Power P, Naas T, Gutkind G. 2021. Redefining the origin and evolution of chromosomally encoded $bla_{CTX-M/KLU}$ in the context of a revised taxonomy of genus *Kluyvera*. Antimicrob Agents Chemother 65. https://doi.org/10.1128/AAC.02424-20

61. Poirel L, Potron A, De La Cuesta C, Cleary T, Nordmann P, Munoz-Price LS. 2012. Wild coastline birds as reservoirs of broad-spectrum-β-lactamase-producing *Enterobacteriaceae* in Miami Beach, Florida. Antimicrob Agents Chemother 56:2756–2758. https://doi.org/10.1128/AAC.05982-11

62. Jamborova I, Janecko N, Halova D, Sedmik J, Mezerova K, Papousek I, Kutilova I, Dolejska M, Cizek A, Literak I. 2018. Molecular characterization of plasmid-mediated AmpC beta-lactamase- and extended-spectrum beta-lactamase-producing *Escherichia coli* and *Klebsiella pneumoniae* among corvids (*Corvus brachyrhynchos* and *Corvus corax*) roosting in Canada. FEMS Microbiol Ecol 94:fiy166. https://doi.org/10.1093/femsec/fiy166

63. Huang YH, Kuan NL, Yeh KS. 2020. Characteristics of extended-spectrum β-lactamase–producing *Escherichia coli* from dogs and cats admitted to a veterinary teaching hospital in Taipei, Taiwan from 2014 to 2017. Front Vet Sci 7:395. https://doi.org/10.3389/fvets.2020.00395

64. Bonnet R. 2004. Growing group of extended-spectrum β-lactamases: the CTX-M enzymes. Antimicrob Agents Chemother 48:1–14. https://doi.org/10.1128/AAC.48.1.1-14.2004

65. Espadamala A, Pereira R, Pallarés P, Lago A, Silva-Del-Río N. 2018. Metritis diagnosis and treatment practices in 45 dairy farms in California. J Dairy Sci 101:9608–9616. https://doi.org/10.3168/jds.2017-14296

66. Hurtado R, Barh D, Weimer BC, Viana MVC, Profeta R, Sousa TJ, Aburjaile FF, Quino W, Souza RP, Mestanza O, Gavilán RG, Azevedo V. 2022. WGS-based lineage and antimicrobial resistance pattern of *Salmonella* Typhimurium isolated during 2000-2017 in Peru. Antibiotics (Basel) 11:1170. https://doi.org/10.3390/antibiotics11091170

67. Tyson GH, McDermott PF, Li C, Chen Y, Tadesse DA, Mukherjee S, Bodeis-Jones S, Kabera C, Gaines SA, Loneragan GH, Edrington TS, Torrence M, Harhay DM, Zhao S. 2015. WGS accurately predicts antimicrobial resistance in *Escherichia coli*. J Antimicrob Chemother 70:2763–2769. https://doi.org/10.1093/jac/dkv186

68. Stubberfield E, AbuOun M, Sayers E, O'Connor HM, Card RM, Anjum MF. 2019. Use of whole genome sequencing of commensal *Escherichia coli* in pigs for antimicrobial resistance surveillance. Euro Surveill 24:1900136. https://doi.org/10.2807/1560-7917.ES.2019.24.50.1900136

69. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, Tyson GH, Zhao S, Hsu C-H, McDermott PF, Tadesse DA, Morales C, Simmons M, Tillman G, Wasilenko J, Folster JP, Klimke W. 2019. Validating the AMRFINder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob Agents Chemother 63:11. https://doi.org/10.1128/AAC.00483-19

70. Messele YE, Alkhallawi M, Veltman T, Trott DJ, McMeniman JP, Kidd SP, Low WY, Petrovski KR. 2022. Phenotypic and genotypic analysis of antimicrobial resistance in *Escherichia coli* recovered from feedlot beef cattle in Australia. Animals (Basel) 12:2256. https://doi.org/10.3390/ani12172256

71. Pascottini OB, Van Schyndel SJ, Spricigo JFW, Rousseau J, Weese JS, LeBlanc SJ. 2020. Dynamics of uterine microbiota in postpartum dairy cows with clinical or subclinical endometritis. Sci Rep 10:12353. https://doi.org/10.1038/s41598-020-69317-z

72. Jeon SJ, Galvão KN. 2018. An advanced understanding of uterine microbial ecology associated with metritis in dairy cows. Genomics Inform 16:e21. https://doi.org/10.5808/GI.2018.16.4.e21

73. Basbas C, Garzon A, Schlesener C, van Heule M, Profeta R, Weimer BC, Silva-Del-Rio N, Byrne BA, Karle B, Aly SS, Lima FS, Pereira RV. 2023. Unveiling the microbiome during post-partum uterine infection: a deep shotgun sequencing approach to characterize the dairy cow uterine microbiome. Anim Microbiome 5:59. https://doi.org/10.1186/s42523-023-00281-5

74. Amos MR, Healey GD, Goldstone RJ, Mahan SM, Düvel A, Schuberth H-J, Sandra O, Zieger P, Dieuzy-Labaye I, Smith DGE, Sheldon IM. 2014. Differential endometrial cell sensitivity to a cholesterol-dependent cytolysin links *Trueperella pyogenes* to uterine disease in cattle. Biol Reprod 90:54. https://doi.org/10.1095/biolreprod.113.115972

75. Rasko DA, Rosovitz MJ, Myers GSA, Mongodin EF, Fricke WF, Gajer P, Crabtree J, Sebaihia M, Thomson NR, Chaudhuri R, Henderson IR, Sperandio V, Ravel J. 2008. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. J Bacteriol 190:6881–6893. https://doi.org/10.1128/JB.00619-08

76. Desvaux M, Dalmasso G, Beyrouthy R, Barnich N, Delmas J, Bonnet R. 2020. Pathogenicity factors of genomic islands in intestinal and extraintestinal *Escherichia coli*. Front Microbiol 11:2065. https://doi.org/10.3389/fmicb.2020.02065

77. Baumgartner M, Zirnbauer R, Schlager S, Mertens D, Gasche N, Sladek B, Herbold C, Bochkareva O, Emelianenko V, Vogelsang H, Lang M, Klotz A, Moik B, Makristathis A, Berry D, Dabsch S, Khare V, Gasche C. 2022. Atypical enteropathogenic *E. coli* are associated with disease activity in ulcerative colitis. Gut Microbes 14:2143218. https://doi.org/10.1080/19490976.2022.2143218

78. Weis AM, Clothier KA, Huang BC, Kong N, Weimer BC. 2016. Draft genome sequences of *Campylobacter jejuni* strains that cause abortion in livestock. Genome Announc 4:e01324-16. https://doi.org/10.1128/genomeA.01324-16

79. Miller JJ, Weimer BC, Timme R, Lüdeke CHM, Pettengill JB, Bandoy DD, Weis AM, Kaufman J, Huang BC, Payne J, Strain E, Jones JL. 2021. Phylogenetic and biogeographic patterns of *Vibrio parahaemolyticus* strains from North America inferred from whole-genome sequence data. Appl Environ Microbiol 87:e01403-20. https://doi.org/10.1128/AEM.01403-20

80. Weimer BC. 2017. 100K pathogen genome project. Genome Announc 5:e00594-17. https://doi.org/10.1128/genomeA.00594-17

81. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

82. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170

83. SeemannT. 2020. Shovill. Available from: https://github.com/tseemann/shovill

84. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114

85. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. Genome Biol 20:257. https://doi.org/10.1186/s13059-019-1891-0

86. Miller JJ, Weimer BC, Timme R, Lüdeke CHM, Pettengill JB, Bandoy DD, Weis AM, Kaufman J, Huang BC, Payne J, Strain E, Jones JL. 2021. Phylogenetic and biogeographic patterns of *Vibrio parahaemolyticus* strains from North America inferred from whole-genome sequence data. Appl Environ Microbiol 87:e01403-20. https://doi.org/10.1128/AEM.01403-20

87. Hernández-Juárez LE, Camorlinga M, Méndez-Tenorio A, Calderón JF, Huang BC, Bandoy DDR, Weimer BC, Torres J. 2021. Analyses of publicly available *Hungatella hathewayi* genomes revealed genetic distances indicating they belong to more than one species. Virulence 12:1950–1964. https://doi.org/10.1080/21505594.2021.1950955

88. Haiminen N, Edlund S, Chambliss D, Kunitomi M, Weimer BC, Ganesan B, Baker R, Markwell P, Davis M, Huang BC, Kong N, Prill RJ, Marlowe CH, Quintanar A, Pierre S, Dubois G, Kaufman JH, Parida L, Beck KL. 2019. Food authentication from shotgun sequencing reads with an application on high protein powders. NPJ Sci Food 3:24. https://doi.org/10.1038/s41538-019-0056-6

89. Bandoy DDR, Weimer BC. 2019. Biological machine learning combined with bacterial population genomics reveals common and rare allelic variants of genes to cause disease. bioRxiv. https://doi.org/10.1101/739540

90. Kaufman JH, Elkins CA, Davis M, Weis AM, Huang BC, Mammel MK, Patel IR, Beck KL, Edlund S, Chambliss D, Douglas J, Bianco S, Kunitomi M, Weimer BC. 2019. Insular microbiogeography: three pathogens as exemplars. In Microbial ecology: current advances from genomics, metagenomics and other omics

91. Titus Brown C, Irber L. 2016. sourmash: a library for MinHash sketching of DNA. J Open Source Softw 1:27. https://doi.org/10.21105/joss.00027

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

93. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693. https://doi.org/10.1093/bioinformatics/btv421

94. Hadfield J, Croucher NJ, Goater RJ, Abudahab K, Aanensen DM, Harris SR. 2018. Phandango: an interactive viewer for bacterial population genomics. Bioinformatics 34:292–293. https://doi.org/10.1093/bioinformatics/btx610

95. Brynildsrud O, Bohlin J, Scheffer L, Eldholm V. 2016. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. Genome Biol 17:238. https://doi.org/10.1186/s13059-016-1108-8

96. SeemannT. 2015. Rapid haploid variant calling and core genome alignment. Available from: https://github.com/tseemann/snippy

97. SeemannT. 2020. Github. Abricate. Available from: https://github.com/tseemann/abricate

98. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, Wlodarski MA, Edalatmand A, Petkau A, Syed SA, Tsang KK, et al. 2023. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the comprehensive antibiotic resistance database. Nucleic Acids Res 51:D690–D699. https://doi.org/10.1093/nar/gkac920

99. Liu B, Zheng D, Zhou S, Chen L, Yang J. 2022. VFDB 2022: a general classification scheme for bacterial virulence factors. Nucleic Acids Res 50:D912–D917. https://doi.org/10.1093/nar/gkab1107

100. Ingle DJ, Valcanis M, Kuzevski A, Tauschek M, Inouye M, Stinear T, Levine MM, Robins-Browne RM, Holt KE. 2016. *In silico* serotyping of *E. coli* from short read data identifies limited novel O-loci but extensive diversity of O:H serotype combinations within and between pathogenic lineages. Microb Genom 2:e000064. https://doi.org/10.1099/mgen.0.000064

101. Oksanen J, Simpson G, Blanchet G, Kindt R, Legendre P, Minchin P, O'Hara RB, Solymos P, Stevens MHH, Szoecs E, et al. 2022. vegan: community ecology package [R package vegan version 2.6-4]. https://CRAN.R-project.org/package=vegan.