Characterizing Plasmids in Bacteria Species Relevant to Urinary Health

Cesar Montelongo Hernandez,[a] Catherine Putonti,[b,c] Alan J. Wolfe[a]

aDepartment of Microbiology and Immunology, Stritch School of Medicine, Loyola University Chicago, Maywood, Illinois, USA
bBioinformatics Program, Loyola University Chicago, Chicago, Illinois, USA
cDepartment of Biology, Loyola University Chicago, Chicago, Illinois, USA

ABSTRACT The urinary tract has a microbial community (the urinary microbiota or urobiota) that has been associated with human health. Whole genome sequencing of bacteria is a powerful tool, allowing investigation of the genomic content of the urobiota, also called the urinary microbiome (urobiome). Bacterial plasmids are a significant component of the urobiome yet are understudied. Because plasmids can be vectors and reservoirs for clinically relevant traits, they are important for urobiota dynamics and thus may have relevance to urinary health. In this project, we sought plasmids in 11 clinically relevant urinary species: *Aerococcus urinae*, *Corynebacterium amycolatum*, *Enterococcus faecalis*, *Escherichia coli*, *Gardnerella vaginalis*, *Klebsiella pneumoniae*, *Lactobacillus gasseri*, *Lactobacillus jensenii*, *Staphylococcus epidermidis*, *Streptococcus anginosus*, and *Streptococcus mitis*. We found evidence of plasmids in *E. faecalis*, *E. coli*, *K. pneumoniae*, *S. epidermidis*, and *S. anginosus* but insufficient evidence in other species sequenced thus far. Some identified plasmidic assemblies were predicted to have putative virulence and/or antibiotic resistance genes, although the majority of their annotated coding regions were of unknown predicted function. In this study, we report on plasmids from urinary species as a first step to understanding the role of plasmids in the bacterial urobiota.

IMPORTANCE The microbial community of the urinary tract (urobiota) has been associated with human health. Whole genome sequencing of bacteria permits examination of urobiota genomes, including plasmids. Because plasmids are vectors and reservoirs for clinically relevant traits, they are important for urobiota dynamics and thus may have relevance to urinary health. Currently, urobiota plasmids are understudied. Here, we sought plasmids in 11 clinically relevant urinary species. We found evidence of plasmids in *E. faecalis*, *E. coli*, *K. pneumoniae*, *S. epidermidis*, and *S. anginosus* but insufficient evidence in the other 6 species. We identified putative virulence and/or antibiotic resistance genes in some of the plasmidic assemblies, but most of their annotated coding regions were of unknown function. This is a first step to understanding the role of plasmids in the bacterial urobiota.

KEYWORDS plasmids, urinary tract, microbiota, incompatibility group, rep, genome, microbial, Gram-negative bacteria, Gram-positive bacteria, urinary tract infection

Until recently, urine was assumed to be sterile (1, 2). It is now confirmed that resident microbial communities (microbiota) exist even in asymptomatic individuals and that specific species are associated with urinary conditions, such as overactive bladder (OAB), urge urinary incontinence (UUI), and urinary tract infection (UTI) (3–6). Multiple bacterial species make up the urinary microbiota (urobiota), with species composition, cell count, and population dynamics linked to urinary conditions or lack of symptoms (7–9). The

bacterial urinary microbiome (urobiome) includes not just the sequence of a bacteria's chromosome but also mobile genetic elements, such as prophages and plasmids (10–12). Plasmids can serve as transmissible reservoirs for clinically relevant traits, such as antibiotic resistance, virulence, and fitness genes (13–15). Plasmids can be genetically heterogenous; therefore, conserved replicon genes such as those involved in replication (*rep*) or incompatibility (*inc*) are used for profiling (16–19). Bacterial plasmids in the urinary tract are understudied yet could be relevant to urinary health and clinical management, as they are in other microbiota niches (20–22). For example, an urgent question in urinary research is why some bacteria may be present in both asymptomatic people and those with lower urinary tract symptoms (LUTS); given their role in pathogenicity, plasmids are a logical target to study in this context. In this project, we analyzed the presence and properties of plasmids in 11 clinically relevant urinary species: *A. urinae*, *C. amycolatum*, *E. faecalis*, *E. coli*, *G. vaginalis*, *K. pneumoniae*, *L. gasseri*, *L. jensenii*, *S. epidermidis*, *S. anginosus*, and *S. mitis*. These commensals, pathogens, and opportunistic pathogens are representative of the diverse taxa commonly detected in the urinary tract (3, 23–31). *A. urinae*, *C. amycolatum*, *E. coli*, *E. faecalis*, *K. pneumoniae*, *S. epidermidis*, *S. anginosus*, and *S. mitis* have been associated with UTI, although all these species can also be present in asymptomatic individuals (3, 23, 24, 27, 29–32). In particular, *E. coli* and *K. pneumoniae* are the species most commonly associated with UTI (33–35). *Gardnerella* species can be present in the bladders of women asymptomatic for urinary conditions, although one study found *Gardnerella* in a larger proportion of participants with OAB (23). *Lactobacillus* species are generally thought to be commensal bacteria, as appears to be the case for *L. jensenii*, which has been shown to have antibacterial activity against *E. coli* (28, 36). In contrast, *L. gasseri* is more frequently detected in women with UUI (23), but the function of this *Lactobacillus* species is unknown.

Current whole genome sequencing (WGS) studies of urinary tract bacteria species have focused on their overall genomic content, with less emphasis on any plasmids they might contain (26, 30, 37–39). *E. coli* is the species most commonly associated with UTI (33–35), and its plasmids in the urinary tract are some of the best-studied relative to other urinary species (40–42). F plasmids, often associated with *incF* loci, are an especially important plasmid type in Enterobacteriaceae, the family that includes both *E. coli* and *K. pneumoniae*. The F plasmid is considered to be widespread and able to frequently transmit both antibiotic resistance and virulence genes (15, 20, 43). IncF plasmids usually are considered to have a narrow host range, only being transferred and maintained in Enterobacteriaceae (44, 45). Plasmids have been identified in in *K. pneumoniae* isolates from patients with UTI and have been proposed as a fingerprint to identify the origin of infection (46, 47).

Plasmids in Gram-positive organisms are significantly less characterized, although progress has been made in clinically relevant bacteria, such as *Staphylococcus* species (17, 48, 49). Staphylococci plasmids are relevant to antibiotic resistance and virulence, and gene content from these can be homologous to other Gram-positive species (17, 49, 50). Plasmids have been profiled in *Staphylococcus saprophyticus* isolates from women with UTI (51, 52). There is much less known about the role that plasmids from other species play in the urinary tract, but we can make inferences from our broader understanding of their plasmid biology. Plasmids in *Enterococcus* are credited for its infamous multi-drug resistance (17, 53). *Lactobacillus* and *Streptococcus* species have evidence of plasmids but are understudied relative to other species mentioned here (54–57). Finally, very little is known of plasmids in *A. urinae* and C. amycolatum, even in the broader microbiology literature (58, 59).

In this study, we utilized genomics to (i) identify plasmids in representative strains of clinically relevant urinary species and (ii) profile plasmid-encoded loci relevant to urinary health. We provide evidence of plasmid presence by generating plasmidic assemblies, identifying replicon-associated loci, and analyzing similarity to the plasmid entries in the NCBI nr/nt database. There is evidence of plasmids in 5 of the 11 urinary species tested (*E. coli*, *E. faecalis*, *K. pneumoniae*, *S. epidermidis*, and *S. anginosus*), some of which carry classic antibiotic resistance and virulence genes. In addition, most

TABLE 1 Replicon, virulence, and antibiotic resistance gene profiling of urinary bacteria plasmidic assemblies

Strain that generated plasmidic assembly	Species	Inc/rep hits	Virulence hits	Antibiotic resistance hits
UMB9250	*Escherichia coli*	IncX1		
UMB1195	*Escherichia coli*	IncB/O/K/Z	traT	
UMB1284	*Escherichia coli*	IncFIA, IncFII	traT	Tetracycline
UMB1284_2	*Escherichia coli*	IncX1		
UMB1180	*Escherichia coli*	Col440I		
UMB7764	*Escherichia coli*	IncFIB, IncFII	sitA, MntB_1, mntB_2, mntB_3	Trimethoprim, sulfamethoxazole, amoxicillin, beta-lactam, multi-drug
UMB9246	*Escherichia coli*	IncFIA, IncFIB	iucC, iutA, sitA, mntB_1, mntB_2, mntB_3	Tetracycline
UMB7780	*Enterococcus faecalis*		cylA, cylL, cylM	
UMB0843	*Enterococcus faecalis*		agg	
UMB7783	*Klebsiella pneumoniae*	IncFII(K)	traT	
UMB8492	*Klebsiella pneumoniae*			
UMB7779	*Klebsiella pneumoniae*	IncFII(K)	traT	
UMB8493	*Staphylococcus epidermidis*			
UMB1227	*Staphylococcus epidermidis*	rep20		Macrolide, bacitracin
UMB0626	*Staphylococcus epidermidis*			Macrolide, antiseptic/disinfectant
UMB0593	*Staphylococcus epidermidis*			
UMB1201	*Staphylococcus epidermidis*			
UMB9183	*Staphylococcus epidermidis*	rep7a		Tetracycline
UMB0567	*Streptococcus anginosus*			
UMB8616	*Streptococcus anginosus*			

putative plasmids have numerous predicted open reading frames that code for proteins with unknown function, underscoring the need for more detailed studies of plasmids in the urinary microbiota. Our findings may be used to prioritize urinary bacterial species whose plasmids could be profiled sufficiently by WGS, while identifying species that require more in-depth analysis, potentially with a wet lab component for plasmid identification. Plasmid research in the urinary tract is an understudied topic, especially in the context of LUTS, and understanding plasmids in these species could provide insight into the role of bacteria in urinary health.

RESULTS

First, we analyzed a set of draft genomes from 71 urinary isolates from 11 different species for presence of plasmids, including species with well-studied plasmids (*E. coli*, *E. faecalis*, *K. pneumoniae*, and *S. epidermidis*) and less well-studied species (*A. urinae*, *C. amycolatum*, *G. vaginalis*, *L. gasseri*, *L. jensenii*, *S. anginosus*, and *S. mitis*). Most strains were isolated from females (N = 69) with or without lower urinary tract symptoms; full details are listed in Table S1 in the supplemental material. To achieve that goal, we generated plasmidic assemblies from the raw reads of these isolates using plasmidSPAdes, which is optimized for identifying and assembling circular sequences, such as plasmids. Plasmidic assemblies could not be generated from the raw sequence reads for any of the *A. urinae*, *G. vaginalis*, and *S. mitis* isolates, and most of the isolates from *C. amycolatum*, *L. gasseri*, and *L. jensenii* (Table S1). Plasmidic assemblies were generated for the rest of the isolates; we searched these for replicon loci via PlasmidFinder. Five urinary species (*E. coli*, *E. faecalis*, *K. pneumoniae*, *S. epidermidis*, *S. anginosus*) were predicted to include replicon (*inc* and *rep*) loci in some of their plasmidic assemblies (Table 1). We used VirulenceFinder and ResFinder to identify virulence and antibiotic resistance genes in the plasmidic assemblies. Most plasmidic assemblies had predicted virulence and/or antibiotic resistance genes (Table 1).

Six of nine *E. coli* isolates had evidence of at least one predicted plasmidic assembly; one of these had evidence of two. We confirmed the bioinformatic prediction of these plasmids for three of the *E. coli* strains by Nanopore long-read sequencing, generating hybrid assemblies for the strains (Table 2). The replicons represented in the urinary *E. coli* plasmidic assemblies were Col440I, IncB/O/K/Z, IncFIA, IncFIB, IncFII, and IncX1

TABLE 2 Comparison of predicted plasmidic assemblies from plasmidSPAdes assembly of short-read sequencing and long-read assembly

Strain	Length of short-read assembled plasmid (topology)	Length of long-read assembled plasmid (topology)
UMB1180	4990 (linear)	4863 (circular)
UMB1195	94010 (linear)	85640 (circular)
UMB1284	75565 (linear)	98469 (linear)
UMB1284_2	35646 (linear)	35519 (circular)

(Table 1). From the seven *E. coli* plasmidic assemblies, there were virulence genes in four and antibiotic resistance genes in three (Table 1). The plasmidic assembly from UMB7764, predicted to be an F plasmid, had three distinct predicted genes for antibiotic resistance against trimethoprim, sulfamethoxazole, and beta-lactams. Plasmidic assemblies from *E. coli* strains UMB1195 and UMB1284 were predicted to encode *traT*, a factor that blocks invasion of similar or same plasmids and protects bacteria against some bacteriophages and killing by animal blood serum (60–63). TraT is thus considered a virulence factor. The plasmidic assembly in UMB7764 had predicted metal scavenging genes (including *sitA*, which can bind metals and aid in adhesion) and manganese transporters (*mntB_1*, *mntB_2*, and *mntB_3*). Lastly, the UMB9246 isolate's plasmidic assembly was predicted to encode *sitA*, as well as the aerobactin synthase *iucC* (involved in iron transport) and *iutA* (which binds aerobactin).

All three putative *K. pneumoniae* plasmids had high query coverage and sequence identity (>90%) to annotated complete plasmid records in GenBank (Table 3). Two *K. pneumoniae* plasmidic assemblies (UMB7783, UMB7779) had the *Klebsiella* variant of the *incFII* gene, which is associated with virulence plasmids (Table 1). UMB7783 and UMB7779 were predicted to encode TraT (Table 1).

No replicon or antibiotic resistance gene was predicted in the three identified *E. faecalis* plasmidic assemblies (Table 1). However, the putative plasmid sequence from *E. faecalis* strain UMB0843 was predicted to encode the virulence gene *agg*, involved in agglutination and adhesion, specifically in the context of promoting content for plasmid conjugation (64). The plasmidic assembly from *E. faecalis* UMB7780 was predicted to encode genes from the cytolysin operon consisting of the activator (*cylA*) and lysin (*cylL*, *cylM*).

For the six plasmidic assemblies of *S. epidermidis*, only UMB1227 (*rep20*) and UMB9183 (*rep7A*) had identifiable replicon loci (Table 1). Macrolide resistance was predicted in UMB1227 and UMB0626, while tetracycline resistance was predicted in UMB9183. The two plasmidic assemblies from *S. anginosus* had no predicted virulence or antibiotic resistance genes.

Plasmidic assemblies were compared to entries in the NCBI nr/nt database via BLAST. Those with relatively high homology (over 20% sequence query coverage) to plasmid entries are listed in Table 2, and those with lower homology (less than 20% sequence query coverage) to plasmid entries are in Table S2; plasmidic assemblies with no hits to plasmid entries are not listed. Relative to known plasmids, five of the seven *E. coli* plasmidic assemblies had a query coverage over 80% and sequence identity over 99%, in addition to a similar sequence length to the reference plasmid entry (Table 3). All three *K. pneumoniae* plasmidic assemblies had high homology to their respective database records and were estimated to be relatively large (~100 kbp) (Table 3). One *E. faecalis* isolate had no evidence of a plasmid (i.e., no plasmidic assembly), but the two other isolates had plasmidic assemblies that had 47% and 77% sequence query coverage, respectively, and over 99% identity to *Enterococcus* plasmids in the database (Table 3). Six of the nine *S. epidermidis* isolates had plasmidic assemblies of 20k–50k bases; although their query coverage was in the 30–50% range, their sequence identity was >98% relative to plasmid entries (Table 3). Two of eight *S. anginosus* isolates had plasmidic assemblies, approximately 50k bases in size, and 32% or 86% sequence query coverage over 97% identity, respectively (Table 3). Some isolates

TABLE 3 Alignment of urinary plasmidic assemblies to reference plasmids

Strain	Taxonomy	Urinary plasmid size (bp)	Plasmid hit	NCBI plasmid size (bp)	Sequence query coverage	E value	Per. Ident	Accession	PMID
UMB9250	*Escherichia coli*	43458	p30155-2	54008	56%	0	98.65%	CP053050.1	32883017
UMB1195	*Escherichia coli*	94010	p86	86147	81%	0	99.18%	CP023387.1	29102123
UMB1284_1	*Escherichia coli*	75565	p179-1	122483	85%	0	100.00%	CP041560.1	32042895
UMB1284_2	*Escherichia coli*	35646	p51008369SK1_C	33826	95%	0	99.95%	CP029976.1	N/A
UMB1180	*Escherichia coli*	4990	pEcI5-3	4863	100%	0	99.98%	CP047739.1	33122675
UMB7764	*Escherichia coli*	58200	p1658/97	125491	69%	0	99.81%	AF550679.1	17220406
UMB9246	*Escherichia coli*	73121	pSCU-313-1	105394	87%	0	99.84%	CP051695.1	32759337
UMB7780	*Enterococcus faecalis*	64875	p26975_2#26	66716	47%	0	99.91%	LR962696.1	N/A
UMB0843	*Enterococcus faecalis*	47683	p26975_1#7	56311	77%	0	98.86%	LR961992.1	N/A
UMB7783	*Klebsiella pneumoniae*	96352	pAR_0096	100759	100%	0	100.00%	CP027614.1	N/A
UMB8492	*Klebsiella pneumoniae*	99217	pAR_0096	100759	100%	0	99.94%	CP027614.1	N/A
UMB7 779	*Klebsiella pneumoniae*	102300	pKpn3-L132	150325	80%	0	99.87%	CP040025.1	31665400
UMB8493	*Staphylococcus epidermidis*	47203	pSE1	51568	42%	0	98.84%	CP066375.1	N/A
UMB1227	*Staphylococcus epidermidis*	49935	pSP01	76991	52%	0	99.87%	KR230047.1	26472766
UMB0626	*Staphylococcus epidermidis*	38439	pFDAARGOS_161	21267	58%	0	99.60%	CP014130.1	N/A
UMB0593	*Staphylococcus epidermidis*	19949	pER01533.3	26684	23%	0	99.90%	CP030674.1	N/A
UMB1201	*Staphylococcus epidermidis*	47834	pSESURV_p1_0612	51026	32%	0	99.95%	CP043786.1	32004459
UMB9183	*Staphylococcus epidermidis*	4566	pSEP1	4439	100%	0	99.98%	AP019722.1	N/A
UMB0567	*Streptococcus anginosus*	6660	pDRPIS7493	4727	32%	0	97.65%	CP002926.1	21994930
UMB8616	*Streptococcus anginosus*	4935	paSTHERMO	4451	86%	0	97.40%	LR822024.1	N/A

from *C. amycolatum* and the *Lactobacillus* species had plasmidic assemblies with low query coverage to plasmid entries in the NCBI nr/nt database (Table S1, Table S2). Raw sequencing reads were mapped to the plasmid sequence in the NCBI database with the highest homology score, which improved the query coverage when comparing urinary isolates to that reference (Table S3).

We annotated and then counted all (open reading frames) ORFs in the plasmidic assemblies; the greatest number of ORFs (N = 241) was in the plasmidic assembly from *E. coli* UMB1195 (~94k bp) (Table 4). We calculated the ratio of ORFs with a predicted function to all predicted ORFs (predicted and hypothetical function). The overall percentage of ORFs annotated with an assigned function ranged from zero to 60.71% in all assemblies, with the highest percentages present in plasmidic assemblies from *S. epidermidis*, *E. coli*, and *K. pneumoniae*. On average, in all the plasmidic assemblies, 31.38% of ORFs were annotated with an assigned function.

Next, we reviewed the plasmidic assembly annotations for genes involved in plasmid biology (transfer, replication, and retention). Only two bacterial species, *E. coli* and *K. pneumoniae*, had known conjugation genes in their plasmidic assemblies. Per the Prokka annotation, these consisted of transfer (*tra*) genes or virB-virD4 genes (Table 5). Genes that block plasmid fertility were predicted in *E. coli* and *K. pneumoniae* plasmidic assemblies, with *finO* genes present in *E. coli* UMB7764, *E. coli* UMB9246, *K. pneumoniae* UMB7783, and *K. pneumoniae* UMB8492. *E. coli* UMB1284 and UMB9246 had a complete module of the toxin-antitoxins (TA) *ccdAB* and *pemIK*, which are involved in plasmid addiction function. *E. coli* UMB9250 and *S. epidermidis* UMB8493 were predicted to code the toxin component of a TA, but the antitoxin gene was not annotated by Prokka.

DISCUSSION

An urgent goal in urobiome research is to elucidate the mechanisms that link the urobiota to urinary conditions (1, 8, 9). This includes understanding why species can be associated with both asymptomatic and symptomatic states (65, 66). Plasmid content of urinary species may be a key component underlying urobiota behavior and their effect on urinary health (20, 67). Plasmids are important reservoirs and vectors for genetic content in bacteria populations, including virulence and antibiotic resistance genes (17, 43, 68). In this study, we analyzed the plasmid content of representative strains of 11 highly relevant

TABLE 4 Summary of annotated content in urinary bacteria plasmidic assemblies

Isolate	Species	Plasmidic assembly size (bp)	Total ORF predicted	ORF annotated with function	ORF annotated as hypothetical	% annotated with function
UMB1180	Escherichia coli	4990	4	0	4	0
UMB1195	Escherichia coli	94010	241	29	212	12.03
UMB1284	Escherichia coli	75565	78	38	40	48.72
UMB1284_2	Escherichia coli	35646	47	11	36	23.4
UMB7764	Escherichia coli	58200	61	35	26	57.38
UMB9246	Escherichia coli	73121	76	40	36	52.63
UMB9250	Escherichia coli	43458	53	15	38	28.3
					avg %	31.78
UMB0843	Enterococcus faecalis	47683	51	6	45	11.76
UMB7780	Enterococcus faecalis	64875	110	37	73	33.63
					avg %	22.7
UMB7779	Klebsiella pneumoniae	102300	108	53	55	49.07
UMB7783	Klebsiella pneumoniae	96352	110	39	71	35.45
UMB8492	Klebsiella pneumoniae	99217	121	40	81	33.06
					avg %	39.19
UMB0567	Streptococcus anginosus	6660	6	0	6	0
UMB8616	Streptococcus anginosus	4935	4	1	3	25
					avg %	12.5
UMB0593	Staphylococcus epidermidis	19949	22	6	16	27.27
UMB0626	Staphylococcus epidermidis	38439	4	1	3	25
UMB1201	Staphylococcus epidermidis	47834	28	17	11	60.71
UMB1227	Staphylococcus epidermidis	49935	44	28	16	63.63
UMB8493	Staphylococcus epidermidis	47203	50	11	39	22
UMB9183	Staphylococcus epidermidis	4566	4	2	2	50
					avg %	41.43
					Total avg %	31.38

urinary species, broadly grouped in three categories: (i) urinary species in which we could consistently detect plasmid presence and profile gene content (*E. coli*, *K. pneumoniae*), (ii) urinary species in which we could detect putative plasmids but most of their content was not profiled via PlasmidFinder or BLAST (*E. faecalis*, *S. anginosus*, *S. epidermidis*), and (ii) urinary species where plasmid presence was negative or inconclusive (*A. urinae*, *C. amycolatum*, *G. vaginalis*, *L. gasseri*, *L. jensenii*, *S. mitis*).

Six of nine urinary *E. coli* isolates had evidence of at least one plasmid, with an average 31.78% of ORFs assigned a known function (Tables 1–4). *E. coli* plasmids are some of the best characterized, with F plasmids being especially relevant due to their prevalence, persistence, and transmission of clinically relevant traits, such as antibiotic resistance and virulence factors (15, 20). There is evidence that F plasmids of urinary *E. coli* are more often found in *E. coli* linked to UTI in kidney-transplanted patients relative to control (41). The *incX1*, *incB/O/K/Z*, and *col4400I* loci were also found in the *E. coli* plasmidic assemblies. IncX and IncB/O/K/Z plasmids have been associated with antibiotic resistance, while colicin plasmids are utilized in competition between similar Gram-negative species (69–72). Broadly speaking, plasmids were readily identified in the urinary *E. coli* isolates either by the presence of replicon loci or homology to plasmid entries in the NCBI nr/nt database. This is likely because of the well-developed plasmid reference databases and assembly algorithms compatible with *E. coli*'s plasmid genetic content (73, 74). Despite these well-developed databases, only about a third of the predicted ORFs were assigned a function.

Profiling of plasmid content in *K. pneumoniae* was successful, likely because of its genetic similarity to *E. coli* (44, 75). The three *K. pneumoniae* putative plasmids had an average 39.19% of ORFs assigned a known function (Tables 1, 3, and 4). All three putative *K. pneumoniae* plasmids had high query coverage and sequence identity (>90%) to annotated complete plasmid records in GenBank (Table 3). The *inc* loci present in *K. pneumoniae* are the *Klebsiella* variants of Col, IncFIB, and IncFII, which bear similarity to *inc* loci in *E. coli* and are respectively linked to colicin and F plasmids (76). As stated

TABLE 5 Summary of plasmid-related genes in urinary bacteria plasmidic assemblies

Strain	Species	Plasmid transfer	Plasmid replication	Plasmid retention
UMB9250	*Escherichia coli*	virB4, ptlE, virB9, virB11, virD4		relE
UMB1195	*Escherichia coli*		parM, ssb (plasmid)	
UMB1284	*Escherichia coli*	traD		ccdB, ccdA, pemI, pemK
UMB1284_2	*Escherichia coli*	virB4, virB8, vir9, vir11		
UMB1180	*Escherichia coli*			
UMB7764	*Escherichia coli*	finO, traD, traI	ssb (plasmid), repB, vapB	
UMB9246	*Escherichia coli*	finO	repB	pemK, pemI, ccdA, ccdB
UMB7780	*Enterococcus faecalis*			
UMB0843	*Enterococcus faecalis*			
UMB7783	*Klebsiella pneumoniae*	traA, traM, tra_I_1, traD_1, finO_1, traN, finO_2, traI_2, traS. traD_2	ssb (plasmid)	
UMB8492	*Klebsiella pneumoniae*	traD, traQ, traN, traC, traV, traA, traM, finO	ssb (plasmid)	
UMB7779	*Klebsiella pneumoniae*	fhO, traI, traD, traQ, traN, traC, traV, traA, traY, traJ, traM		
UMB8493	*Staphylococcus epidermidis*			yoeB
UMB1227	*Staphylococcus epidermidis*			
UMB0626	*Staphylococcus epidermidis*			
UMB0593	*Staphylococcus epidermidis*			
UMB1201	*Staphylococcus epidermidis*			
UMB9183	*Staphylococcus epidermidis*			
UMB0567	*Streptococcus anginosus*			
UMB8616	*Streptococcus anginosus*		pre	

before, F plasmids are clinically relevant due to their antibiotic resistant and virulent genetic content (20). The presence of colicin plasmids could point to competition between *K. pneumoniae* and other bacteria species, including *E. coli* (71, 72). Antibiotic resistance genes were not predicted in *K. pneumoniae* plasmidic assemblies, but the virulence gene *traT* was predicted in two of them (Table 1). The plasmidic assembly from UMB7779 had 80% sequence query coverage and 100% identity to the plasmid pKpn3-L132 (CP027614.1) present in *K. pneumoniae* OXA-48, which contributed to a nosocomial outbreak in Taiwan (75).

Plasmids were predicted in six of nine *S. epidermidis* isolates, with an average 41.43% of ORFs assigned a known function (the highest species average) (Tables 1, 3, and 4). Most of these plasmidic assemblies had high sequence identity but low query coverage with their closest plasmid homolog in GenBank (Table 3). Antibiotic resistance was predicted in these plasmidic assemblies (Table 1), which could make them clinically relevant (50). As an example, the plasmidic assembly from UMB1227 has homology to a conjugative plasmid in clinical *S. epidermidis* from Italy with multiple resistances (77).

E. faecalis and *S. anginosus* had plasmidic assemblies that matched plasmid entries in the NCBI nr/nt database, but the tools we employed did not profile genes to the same extent as the three aforementioned species (Tables 1 and 3). The two *E. faecalis* plasmidic assemblies had an average 22.7% of ORFs assigned a known function (Table 4). Neither *E. faecalis* plasmidic assembly had evidence of antibiotic resistance genes, but virulence genes were predicted (Table 1). The two *S. anginosus* isolates with plasmidic assemblies, UMB0567 and UMB8616, had relatively small putative plasmid sequences (5k–6k bp) with only a small number of ORFs predicted and annotated with a function (Tables 1, 3, and 4). These plasmidic assemblies did not have predicted genes for any other plasmid content profiled. The evidence that these isolates have a plasmid relies on their plasmidic assemblies being similar to pDRPIS7493 and paSTHERMO (over 99% identity and respectively 32% and 85% sequence query coverage) (Table 3). However, as *rep* loci were not detected, they may be novel, or these loci were not included in the assembly. Given the lack of replicon loci and relatively low sequence query coverage when compared to known plasmids, further studies are needed to verify that *E. faecalis* UMB7780 and *S. epidermidis* UMB0567 do contain a plasmid.

plasmidSPAdes did not produce a plasmidic assembly for any of the urinary isolates considered here for the species *A. urinae*, *G. vaginalis*, and *S. mitis* (Table S1). The remaining urinary species (*C. amycolatum*, *L. gasseri*, *L. jensenii*) had no convincing evidence of plasmid presence, given the absence of replicon loci and very low homology to known plasmids (Table S1 and S2). The short plasmid-like sequences in *C. amycolatum*, *L. gasseri*, and *L. jensenii* could be due to plasmid-like regions in the chromosome of these isolates (e.g., prophage, past genetic exchange with plasmids). However, the absence of evidence is not evidence of absence. There are few studies on the plasmids of these species, which means few references in the databases for comparison (73, 78, 79) and other sequencing methods are required to identify novel plasmids, for example hybrid assembly. Another factor to consider is the relatively small sample size per species analyzed.

We utilized homologous plasmids in Table 3 as a reference to map raw sequence reads from the respective urinary isolate (Table S3). In most cases, a plasmidic assembly was produced that when compared to its reference had a sequence query coverage and identity over 90%. This provides evidence that the plasmid genetic content in these urinary isolates is similar to known plasmid sequences. However, there is a drawback in relying on this method for plasmid assembly, as it limits output to what is already known. Urinary plasmids may contain novel genetic content that is not present in the plasmid reference sequence.

Another important element to consider in plasmid research is the ability of certain plasmids to transfer within and between species (14). Plasmid conjugation genes were identified in plasmidic assemblies from *E. coli* and *K. pneumoniae*, specifically *tra* and *virB-virD* genes (Table 5) (80, 81). Potentially relevant is that some plasmidic assemblies from *E. coli* are from the same Inc groups as those in *K. pneumoniae* (specifically loci found in F and colicin plasmids) (42, 71). Plasmid exchange is known to occur in *E. coli* and other Gram-negative species, including *Klebsiella*, and exchange could be occurring in the urinary tract (44, 81). *traT* also was identified, which aids plasmids in blocking invasion by similar plasmids, suggesting that plasmid competition occurs in the urinary tract for these species (60, 63) (Table 1). Conjugation genes were not identified in *E. faecalis*, *S. epidermidis*, or *S. anginosus*, which begs the following question: Is conjugation simply not present? (77, 82) Alternatively, do these species utilize a system absent in the gene reference database or one that has not been assigned that functionality? Outside the urinary tract, there is evidence that Enterococci can conjugate plasmids, while species of Staphylococci may also rely on transformation and vertical transmission in addition to conjugation (17, 82, 83). Horizontal exchange of plasmids in these urinary species may necessitate the creation of custom reference databases with conjugation and competence genes for appropriate gene profiling.

In this project, we relied on two types of tools: (i) reference-independent (*de novo*) assembly of plasmidic raw sequencing reads based on algorithmic assumptions (i.e., plasmidSPAdes assembly), and (ii) reference-dependent gene profiling (web BLAST, PlasmidFinder, ResFinder, VirulenceFinder, Prokka) (73, 78, 79, 84–86). For plasmidSPAdes, the assembly algorithm has been reported to assemble plasmids in Enterobacteriaceae (*E. coli*, *K. pneumoniae*) and Staphylococcaceae (*S. epidermidis*) (74); employing long-read sequencing for 3 of the *E. coli* isolates confirmed the capabilities of the plasmidSPAdes tool to properly assemble and identify *E. coli* plasmids (Table 2). We anticipated that plasmid assembly would occur in *E. faecalis*, *S. anginosus*, and *S. mitis* given their genetic overlap with Staphylococci. More unpredictable was plasmid assembly in the other species (*A. urinae*, *C. amycolatum*, *G. vaginalis*, *L. gasseri*, *L. jensenii*), where less is known of their plasmid size, circular/linear composition, and copy number, especially in the urinary tract (54, 58, 59). Consequently, it is in the latter species where we saw either no or inconclusive plasmid results. The results of the reference-dependent method were more predictable, however. Reference-based profiling relies on a robust database so that queries can have matches (84, 86). The databases we utilized are primarily composed of well-studied organisms, such as *E. coli* and *Staphylococcus* species (78, 79, 85). In

urinary research, species such as those in the genera *Aerococcus*, *Corynebacterium*, *Gardnerella*, and *Lactobacillus* are understudied, and much of their genetic content is unknown (23, 26).

A limitation of this study was the use of plasmidSPAdes plasmidic assemblies as representatives of plasmids in these urinary species. While not perfectly accurate, plasmidic assemblies are still a valid representation of plasmid content in bacteria (73, 74). Three of the four plasmidic assemblies produced by plasmidSPAdes were able to be confirmed and closed via long-read sequencing; the fourth, one of the UMB1284 short-read plasmidic assemblies, was identified by long-read sequencing, but the long-read contig was significantly longer than the one produced by plasmidSPAdes (Table 2). Further investigation into this particular strain is needed to ascertain the structure of this plasmid. Initial studies, such as this one, are necessary to build a more complete understanding of plasmids in the urinary tract.

In terms of future projects that study plasmid content in urinary species, we envision two broad strategies. For well-studied bacteria like those from the families Enterobacteriaceae and Staphylococcaceae, it is feasible to organize high-throughput genomic profiling, even from existing WGS raw short reads. This can be streamlined by scripting a pipeline of commonly used assemblers (e.g., plasmidSPAdes) then scanning the output with reference databases (e.g., Prokka, local BLAST) (73, 84, 86). While many urinary *E. coli* and *K. pneumoniae* strains are publicly available, we included just a few here serving as a control, given the wealth of information available for their plasmids and the fact that many tools/databases for plasmid bioinformatics have been benchmarked using *E. coli*. Bacteria from the families Enterococcaceae and Streptococcaceae face more challenges in terms of specific gene profiling, but it is realistic to pinpoint whether plasmids are present in these species with the approach we employed, especially if custom reference databases are utilized. Finally, for less studied bacterial genera (e.g., *Aerococcus*, *Corynebacterium*, *Gardnerella*, *Lactobacillus*), a more arduous approach is required for predicting plasmid presence and profiling genetic content. Our recommendation would be to rely on known plasmid sequences in members of these species isolated from outside the urinary tract to predict presence of plasmids in urinary isolates. Using this strategy, the impetus would be on building reference databases specific to a family or genus. Long-read sequencing of these genera, both from the urobiome and elsewhere, may aid in building this reference database.

In this study, we employed genomics to analyze the presence of plasmids in urinary isolates of 11 bacteria species relevant to urinary health. The contribution of plasmids to urinary health have not been thoroughly assessed, even in species in the families Enterobacteriaceae and Staphylococcaceae, which are arguably the best studied bacteria in microbiology (1, 7, 9). Plasmids have immense clinical and biological relevance, able to retain and transmit key traits like antibiotic resistance and virulence (15, 20). Plasmid research could shed light into the important question of why isolates from many of these species can be present in people that can be both symptomatic and asymptomatic for urinary conditions (65, 66). This study provides a starting point for plasmid content in urobiota species linked to LUTS and brings attention to the need for wet lab research and specialized bioinformatic tools to further characterize bacterial plasmids in the urinary tract.

MATERIALS AND METHODS

The draft genomes of 71 urinary isolates from 11 species (*A. urinae* (N = 7), *C. amycolatum* (N = 5), *E. coli* (N = 9), *E. faecalis* (N = 3), *G. vaginalis* (N = 5), *K. pneumoniae* (N = 3), *L. gasseri* (N = 6), *L. jensenii* (N = 8), *S. epidermidis* (N = 9), *S. anginosus* (N = 8), and *S. mitis* (N = 8)) were previously sequenced, assembled, and made publicly available by our group (BioProject PRJNA316969) (Table S1). To create plasmid assemblies, raw sequence reads (SRA) were downloaded for these species and assembled using plasmidspades.py of SPAdes v3.12 with k values of 55,77,99,127 and the only-assembler parameter (73). plasmidSPAdes takes the assembly graph from SPAdes and classifies a subgraph as the plasmid graph, which is further processed into plasmidic assemblies (87, 88). "Plasmidic assemblies contain plasmid-like sequences from the WGS, though further curation and pruning may be necessary to remove false positives" (73, 74). In this study we assess plasmidic assemblies as a representative of plasmid content. Assemblies were renamed via a Bash script, and contigs less than 500 bp were removed via bioawk. Each contig in the plasmidic assembly was queried via megablast against the nr/nt database, and

contigs with homology to plasmid records were retained for further analysis while contigs with chromosomal homology were removed from further consideration (84). Assemblies were then concatenated as a single read and queried via megablast against the nr/nt database; assemblies were organized on their sequence query coverage, percent identity, and E-value to plasmid entries. The Bowtie2 (version 2.3.2.) plug-in in Geneious Prime v2021.1 was used to verify even coverage of raw sequence reads to the NCBI plasmid record with highest homology to a given plasmid assembly (89). Curated plasmidic assemblies are publicly available through the BioProject and their strain's respective BioSample Assembly (Assembly Database); the BioSample accession numbers for each strain examined are included in Table S1. Furthermore, plasmidic assembly sequences can be directly accessed via FigShare (https://doi.org/10.6084/m9.figshare.17005318.v1).

To identity *rep* and *inc* loci, the putative plasmid assemblies were scanned with PlasmidFinder v2.1, using either the Enterobacteriaceae or Gram-positive database, with a threshold of 95% identity and a minimum 60% coverage (85). To identify known antibiotic resistance genes, the FASTA files were scanned with ResFinder v4.1 using the "acquired antimicrobial resistance genes" option (78). To identify known virulence genes, the FASTA files were scanned with VirulenceFinder v2.0 with an identity threshold of 90%, and the "minimum sequence length of 60%" option (79). Plasmid assemblies were annotated using Prokka v1.14.5 with default parameters in addition to parameters -entre XXX and -compliant (86). Annotation output files were renamed and reorganized using a Bash script. The annotated ORFs were reviewed for predicted functions related to plasmid transfer, replication, and addiction/retention.

E. coli isolates UMB1180, UMB1195, and UMB1284 were sequenced using Nanopore long-reads sequencing. Each strain was streaked from cultures stored at −80°C onto colistin nalidixic acid (CAN) plates and incubated overnight at 35°C with 5% CO_2. Single colonies were selected and grown in liquid LB overnight at 37°C, with shaking. DNA was extracted using the QIAmp DNA minikit with the following exceptions: samples were not vortexed, rather agitated by hand, and the lysis step was conducted for 1.5–2 h at 37°C. DNA was then shipped to the Columbia University Core Facility (New York, NY), where libraries were prepared using the Rapid Barcoding 96 kit (Oxford Nanopore, SQK-RBK110.96) according to the manufacturer's instructions and sequenced on an Oxford Nanopore GridION using an R9.4.2 flow cell. High accuracy basecalling and demultiplexing was performed using MinKNOW v21.05.20. Read QC was performed using Porechop v0.2.4 (https://github.com/rrwick/Porechop) to remove adapter and barcode sequences and mothur v1.25.0 (https://mothur.org/) to remove long homopolymeric regions (>20 bp) and short reads (<1,000 bp). These filtered reads were then assembled using Unicycler v0.4.9 with default parameters (90). Nanopore reads have been deposited in SRA, associated with the BioSample Accession Numbers for the three strains (indicated in Table S1).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.2 MB.

REFERENCES

1. Brubaker L, Wolfe AJ. 2015. The new world of the urinary microbiota in women. Am J Obstet Gynecol 213:644–649. https://doi.org/10.1016/j.ajog.2015.05.032.

2. Wolfe AJ, Brubaker L. 2015. "Sterile urine" and the presence of bacteria. Eur Urol 68:173–174. https://doi.org/10.1016/j.eururo.2015.02.041.

3. Hilt EE, McKinley K, Pearce MM, Rosenfeld AB, Zilliox MJ, Mueller ER, Brubaker L, Gai X, Wolfe AJ, Schreckenberger PC. 2014. Urine is not sterile: use of enhanced urine culture techniques to detect resident bacterial flora in the adult female bladder. J Clin Microbiol 52:871–876. https://doi.org/10.1128/JCM.02876-13.

4. Brubaker L, Nager CW, Richter HE, Visco A, Nygaard I, Barber MD, Schaffer J, Meikle S, Wallace D, Shibata N, Wolfe AJ. 2014. Urinary bacteria in adult women with urgency urinary incontinence. Int Urogynecol J 25:1179–1184. https://doi.org/10.1007/s00192-013-2325-2.

5. Price TK, Dune T, Hilt EE, Thomas-White KJ, Kliethermes S, Brincat C, Brubaker L, Wolfe AJ, Mueller ER, Schreckenberger PC. 2016. The clinical urine culture: enhanced techniques improve detection of clinically relevant microorganisms. J Clin Microbiol 54:1216–1222. https://doi.org/10.1128/JCM.00044-16.

6. Hartmann KE, McPheeters ML, Biller DH, Ward RM, McKoy JN, Jerome RN, Micucci SR, Meints L, fisher JA, Scott TA, Slaughter JC, Blume JD. 2009. Treatment of overactive bladder in women. Evid Rep Technol Assess (Full Rep) no. 187: 09-E017.

7. Wolfe AJ, Brubaker L. 2019. Urobiome updates: advances in urinary microbiome research. Nat Rev Urol 16:73–74. https://doi.org/10.1038/s41585-018-0127-5.

8. Price TK, Wolff B, Halverson T et al. 2020. Temporal dynamics of the adult female lower urinary tract microbiota. mBio 11:e00475-20. https://doi.org/10.1128/mbio.00475-20.

9. Thomas-White K, Brady M, Wolfe AJ, Mueller ER. 2016. The bladder is not sterile: history and current discoveries on the urinary microbiome. Curr Bladder Dysfunct Rep 11:18–24. https://doi.org/10.1007/s11884-016-0345-8.

10. Żaczek M, Weber-Dąbrowska B, Międzybrodzki R, Górski A. 2020. Phage prevalence in the human urinary tract—current knowledge and therapeutic implications. Microorganisms 8:1802. https://doi.org/10.3390/microorganisms8111802.

11. Miller-Ensminger T, Garretto A, Brenner J, Thomas-White K, Zambom A, Wolfe AJ, Putonti C. 2018. Bacteriophages of the urinary microbiome. J Bacteriol 200:e00738-17. https://doi.org/10.1128/JB.00738-17.

12. Garretto A, Miller-Ensminger T, Wolfe AJ, Putonti C. 2019. Bacteriophages of the lower urinary tract. Nat Rev Urol 16:422–432. https://doi.org/10.1038/s41585-019-0192-4.

13. Hughes C, Bauer E, Roberts AP. 1981. Spread of R-plasmids among Escherichia coli causing urinary tract infections. Antimicrob Agents Chemother 20:496–502. https://doi.org/10.1128/AAC.20.4.496.

14. Smillie C, Garcillan-Barcia MP, Francia MV, Rocha EPC, de la Cruz F. 2010. Mobility of plasmids. Microbiol Mol Biol Rev 74:434–452. https://doi.org/10.1128/MMBR.00020-10.

15. Koraimann G. 2018. Spread and persistence of virulence and antibiotic resistance genes: a ride on the F plasmid conjugation module. EcoSal Plus 8(1). https://doi.org/10.1128/ecosalplus.ESP-0003-2018.

16. Wagner S, Lupolova N, Gally DL, Argyle SA. 2017. Convergence of plasmid architectures drives emergence of multi-drug resistance in a clonally diverse Escherichia coli population from a veterinary clinical care setting. Vet Microbiol 211:6–14. https://doi.org/10.1016/j.vetmic.2017.09.016.

17. Jensen LB, Garcia-Migura L, Valenzuela AJS, Løhr M, Hasman H, Aarestrup FM. 2010. A classification system for plasmids from enterococci and other Gram-positive bacteria. J Microbiol Methods 80:25–43. https://doi.org/10.1016/j.mimet.2009.10.012.

18. Shintani M, Sanchez ZK, Kimbara K. 2015. Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. Front Microbiol 6:242. https://doi.org/10.3389/fmicb.2015.00242.

19. Osborn AM, da Silva Tatley FM, Steyn LM, Pickup RW, Saunders JR. 2000. Mosaic plasmids and mosaic replicons: evolutionary lessons from the analysis of genetic diversity in IncFII-related replicons. Microbiology 146:2267–2275. https://doi.org/10.1099/00221287-146-9-2267.

20. Stephens C, Arismendi T, Wright M, Hartman A, Gonzalez A, Gill M, Pandori M, Hess D. 2020. F plasmids are the major carriers of antibiotic resistance genes in human-associated commensal Escherichia coli. mSphere 5:e00709-20. https://doi.org/10.1128/mSphere.00709-20.

21. Stalder T, Press MO, Sullivan S, Liachko I, Top EM. 2019. Linking the resistome and plasmidome to the microbiome. ISME J 13:2437–2446. https://doi.org/10.1038/s41396-019-0446-4.

22. Tepekule B, Zur Wiesch PA, Kouyos RD, Bonhoeffer S. 2019. Quantifying the impact of treatment history on plasmid-mediated resistance evolution in human gut microbiota. Proc Natl Acad Sci U S A 116:23106–23116. https://doi.org/10.1073/pnas.1912188116.

23. Pearce MM, Hilt EE, Rosenfeld AB, Zilliox MJ, Thomas-White K, Fok C, Kliethermes S, Schreckenberger PC, Brubaker L, Gai X, Wolfe AJ. 2014. The female urinary microbiome: a comparison of women with and without urgency urinary incontinence. mBio 5:1283–1297. https://doi.org/10.1128/mBio.01283-14.

24. Cai T, Mazzoli S, Mondaini N, Meacci F, Nesi G, D'Elia C, Malossini G, Boddi V, Bartoletti R. 2012. The role of asymptomatic bacteriuria in young women with recurrent urinary tract infections: to treat or not to treat? Clin Infect Dis 55:771–777. https://doi.org/10.1093/cid/cis534.

25. Alteri CJ, Mobley HLT. 2015. Metabolism and fitness of urinary tract pathogens. Microbiol Spectr 3(3). https://doi.org/10.1128/microbiolspec.MBP-0016-2015.

26. Malki K, Shapiro JW, Price TK, Hilt EE, Thomas-White K, Sircar T, Rosenfeld AB, Kuffel G, Zilliox MJ, Wolfe AJ, Putonti C. 2016. Genomes of Gardnerella strains reveal an abundance of prophages within the bladder microbiome. PLoS One 11:e0166757. https://doi.org/10.1371/journal.pone.0166757.

27. Hyun M, Lee JY, Ah Kim H, Ryu SY. 2019. Comparison of Escherichia coli and Klebsiella pneumoniae acute pyelonephritis in Korean patients. Infect Chemother 51:130–141. https://doi.org/10.3947/ic.2019.51.2.130.

28. Kalyoussef S, Nieves E, Dinerman E, Carpenter C, Shankar V, Oh J, Burd B, Angeletti RH, Buckheit KW, Fredricks DN, Madan RP, Keller MJ, Herold BC. 2012. Lactobacillus proteins are associated with the bactericidal activity against E. coli of female genital tract secretions. PLoS One 7:e49506. https://doi.org/10.1371/journal.pone.0049506.

29. Upadhyayula S, Kambalapalli M, Asmar BI. 2012. Staphylococcus epidermidis urinary tract infection in an infant. Case Rep Infect Dis 2012:983153. https://doi.org/10.1155/2012/983153.

30. Thomas-White K, Forster SC, Kumar N, Van Kuiken M, Putonti C, Stares MD, Hilt EE, Price TK, Wolfe AJ, Lawley TD. 2018. Culturing of female bladder bacteria reveals an interconnected urogenital microbiota. Nat Commun 9:1557. https://doi.org/10.1038/s41467-018-03968-5.

31. Shahinas D, Tamber GS, Arya G, Wong A, Lau R, Jamieson F, Ma JH, Alexander DC, Low DE, Pillai DR. 2011. Whole-genome sequence of Streptococcus pseudopneumoniae isolate IS7493. J Bacteriol 193:6102–6103. https://doi.org/10.1128/JB.06075-11.

32. Sadeghi-Bojd S, Naghshizadian R, Mazaheri M, Ghane Sharbaf F, Assadi F. 2019. Efficacy of probiotic prophylaxis after the first febrile urinary tract infection in children with normal urinary tracts. J Pediatric Infect Dis Soc 9:305–310. https://doi.org/10.1093/jpids/piz025.

33. Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. 2015. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. Nat Rev Microbiol 13:269–284. https://doi.org/10.1038/nrmicro3432.

34. Ronald A. 2002. The etiology of urinary tract infection: traditional and emerging pathogens. Am J Med 113:14–19. https://doi.org/10.1016/S0002-9343(02)01055-0.

35. Foxman B. 2014. Urinary tract infection syndromes: occurrence, recurrence, bacteriology, risk factors, and disease burden. Infect Dis Clin North Am 28:1–13. https://doi.org/10.1016/j.idc.2013.09.003.

36. Mores CR, Price TK, Wolff B, et al. 2021. Genomic relatedness and clinical significance of Streptococcus mitis strains isolated from the urogenital tract of sexual partners. Microb Genom 7:e000535. https://doi.org/10.1099/mgen.0.000535.

37. Garretto A, Thomas-White K, Wolfe AJ, Putonti C. 2018. Detecting viral genomes in the female urinary microbiome. J Gen Virol 99:1141–1146. https://doi.org/10.1099/jgv.0.001097.

38. Price TK, Mehrtash A, Kalesinskas L, Malki K, Hilt EE, Putonti C, Wolfe AJ. 2016. Genome sequences and annotation of two urinary isolates of E. coli. Stand Genomic Sci 11:79. https://doi.org/10.1186/s40793-016-0202-6.

39. Garretto A, Miller-Ensminger T, Ene A, Merchant Z, Shah A, Gerodias A, Biancofiori A, Canchola S, Canchola S, Castillo E, Chowdhury T, Gandhi N, Hamilton S, Hatton K, Hyder S, Krull K, Lagios D, Lam T, Mitchell K, Mortensen C, Murphy A, Richburg J, Rokas M, Ryclik S, Sulit P, Szwajnos T, Widuch M, Willis J, Woloszyn M, Brassil B, Johnson G, Mormando R, Maskeri L, Batrich M, Stark N, Shapiro JW, Montelongo Hernandez C, Banerjee S, Wolfe AJ, Putonti C. 2020. Genomic survey of E. coli from the bladders of women with and without lower urinary tract symptoms. Front Microbiol Sect Infect Dis 11:2094. https://doi.org/10.3389/fmicb.2020.02094.

40. Brolund A, Franzén O, Melefors Ö, Tegmark-Wisell K, Sandegren L. 2013. Plasmidome-analysis of ESBL-producing Escherichia coli using conventional typing and high-throughput sequencing. PLoS One 8:e65793. https://doi.org/10.1371/journal.pone.0065793.

41. Abo Basha J, Kiel M, Görlich D, Schütte-Nütgen K, Witten A, Pavenstädt H, Kahl BC, Dobrindt U, Reuter S. 2019. Phenotypic and genotypic characterization of Escherichia coli causing urinary tract infections in kidney-transplanted patients. JCM 8:988. https://doi.org/10.3390/jcm8070988.

42. Yang Q-E, Sun J, Li L, Deng H, Liu B-T, Fang L-X, Liao X-P, Liu Y-H. 2015. IncF plasmid diversity in multi-drug resistant Escherichia coli strains from animals in China. Front Microbiol 6:964. https://doi.org/10.3389/fmicb.2015.00964.

43. Fernandez-Lopez R, de Toro M, Moncalian G, Garcillan-Barcia MP, de la Cruz F. 2016. Comparative genomics of the conjugation region of F-like plasmids: five shades of F. Front Mol Biosci 3:71. https://doi.org/10.3389/fmolb.2016.00071.

44. Goren MG, Carmeli Y, Schwaber MJ, Chmelnitsky I, Schechner V, Navon-Venezia S. 2010. Transfer of carbapenem-resistant plasmid from Klebsiella pneumoniae ST258 to Escherichia coli in patient. Emerg Infect Dis 16:1014–1017. https://doi.org/10.3201/eid1606.091671.

45. Schjørring S, Struve C, Krogfelt KA. 2008. Transfer of antimicrobial resistance plasmids from Klebsiella pneumoniae to Escherichia coli in the mouse intestine. J Antimicrob Chemother 62:1086–1093. https://doi.org/10.1093/jac/dkn323.

46. Caneiras C, Lito L, Melo-Cristino J, Duarte A. 2019. Community- and hospital-acquired Klebsiella pneumoniae urinary tract infections in Portugal: virulence and antibiotic resistance. Microorganisms 7:138. https://doi.org/10.3390/microorganisms7050138.

47. Donovan WH, Hull R, Cifu DX, Brown HD, Smith NJ. 1990. Use of plasmid analysis to determine the source of bacterial invasion of the urinary tract. Paraplegia 28:573–582. https://doi.org/10.1038/sc.1990.77.

48. Gillaspy AF, Iandolo JJ, Tang YW, Stratton CW. 2019. Staphylococcus, p 309–320. In Schaechter M (ed), Encyclopedia of microbiology. Elsevier, Amsterdam, Netherlands. https://doi.org/10.1016/B978-0-12-801238-3.02304-7.

49. Shahkarami F, Rashki A, Ghalehnoo ZR. 2014. Microbial susceptibility and plasmid profiles of methicillin-resistant Staphylococcus aureus and methicillin-susceptible S. aureus. Jundishapur J Microbiol 7:e16984. https://doi.org/10.5812/jjm.16984.

50. McCarthy AJ, Lindsay JA. 2012. The distribution of plasmids that carry virulence and resistance genes in Staphylococcus aureus is lineage associated. BMC Microbiol 12:104. https://doi.org/10.1186/1471-2180-12-104.

51. Hedman P, Ringertz O, Olsson K, Wollin R. 1991. Plasmid-identified Staphylococcus saprophyticus isolated from the rectum of patients with urinary tract infections. Scand J Infect Dis 23:569–572. https://doi.org/10.3109/00365549109105179.

52. Cristino JA, Pereira AT, Andrade LG. 1989. Diversity of plasmids in *Staphylococcus saprophyticus* isolated from urinary tract infections in women. Epidemiol Infect 102:413–419. https://doi.org/10.1017/s0950268800030120.

53. Arredondo-Alonso S, Top J, McNally A, Puranen S, Pesonen M, Pensar J, Marttinen P, Braat JC, Rogers MRC, van Schaik W, Kaski S, Willems RJL, Corander J, Schürch AC. 2020. Plasmids shaped the recent emergence of the major nosocomial pathogen *Enterococcus faecium*. mBio 11. https://doi.org/10.1128/mBio.03284-19.

54. Wang TT, Lee BH. 1997. Plasmids in Lactobacillus. Crit Rev Biotechnol 17:227–272. https://doi.org/10.3109/07388559709146615.

55. Fang F, Flynn S, Li Y, Claesson MJ, van Pijkeren J-P, Collins JK, van Sinderen D, O'Toole PW. 2008. Characterization of endogenous plasmids from *Lactobacillus salivarius* UCC118. Appl Environ Microbiol 74:3216–3228. https://doi.org/10.1128/AEM.02631-07.

56. Clewell DB. 1981. Plasmids, drug resistance, and gene transfer in the genus Streptococcus. Microbiol Rev 45:409–436. https://doi.org/10.1128/mr.45.3.409-436.1981.

57. Bergmann R, Nerlich A, Chhatwal GS, Nitsche-Schmitz DP. 2014. Distribution of small native plasmids in *Streptococcus pyogenes* in India. Int J Med Microbiol 304:370–378. https://doi.org/10.1016/j.ijmm.2013.12.001.

58. Deb J, Nath N. 1999. Plasmids of corynebacteria. FEMS Microbiol Lett 175:11–20. https://doi.org/10.1111/j.1574-6968.1999.tb13596.x.

59. Yang Q, Zhu Y, Schwarz S, Wang L, Liu W, Yang W, Luan T, Liu S, Zhang W. 2021. A novel plasmid from *Aerococcus urinaeequi* of porcine origin co-harboring the tetracycline resistance genes *tet*(58) and *tet*(61). Vet Microbiol 257:109065. https://doi.org/10.1016/j.vetmic.2021.109065.

60. Sukupolvi S, O'Connor CD. 1990. TraT lipoprotein, a plasmid-specified mediator of interactions between Gram-negative bacteria and their environment. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC372785/pdf/microrev00039-0011.pdf. Accessed 17 June 2019.

61. Pramoonjago P, Kaneko M, Kinoshita T, Ohtsubo E, Takeda J, Hong KS, Inagi R, Inoue K. 1992. Role of TraT protein, an anticomplementary protein produced in Escherichia coli by R100 factor, in serum resistance. J Immunol 148:827-36.

62. Riede I, Eschbach M-L. 1986. Evidence that TraT interacts with OmpA of *Escherichia coli*. FEBS Letters 205:241–245. https://doi.org/10.1016/0014-5793(86)80905-X.

63. Achtman M, Kennedy N, Skurray R. 1977. Cell-cell interactions in conjugating *Escherichia coli*: Role of traT protein in surface exclusion. Proc Natl Acad Sci U S A 74:5104–5108. https://doi.org/10.1073/pnas.74.11.5104.

64. Strateva T, Atanasova D, Savov E, Petrova G, Mitov I. 2016. Incidence of virulence determinants in clinical *Enterococcus faecalis* and *Enterococcus faecium* isolates collected in Bulgaria. Braz J Infect Dis 20:127–133. https://doi.org/10.1016/j.bjid.2015.11.011.

65. Brubaker L, Wolfe AJ. 2017. The female urinary microbiota, urinary health and common urinary disorders. Ann Transl Med 5:34. https://doi.org/10.21037/atm.2016.11.62.

66. Brubaker L, Wolfe A. 2016. The urinary microbiota: a paradigm shift for bladder disorders? Curr Opin Obstet Gynecol 28:407–412. https://doi.org/10.1097/GCO.0000000000000298.

67. Froggatt JW, Johnston JL, Galetto DW, Archer GL. 1989. Antimicrobial resistance in nosocomial isolates of *Staphylococcus haemolyticus*. Antimicrob Agents Chemother 33:460–466. https://doi.org/10.1128/AAC.33.4.460.

68. Carattoli A. 2009. Resistance plasmid families in *Enterobacteriaceae*. Antimicrob Agents Chemother 53:2227–2238. https://doi.org/10.1128/AAC.01707-08.

69. Ma Y, Chen J, Fong K, Nadya S, Allen K, Laing C, Ziebell K, Topp E, Carroll LM, Wiedmann M, Delaquis P, Wang S. 2021. Antibiotic resistance in shiga toxigenic *Escherichia coli* isolates from surface waters and sediments in a mixed use urban agricultural landscape. Antibiotics 10:237. https://doi.org/10.3390/antibiotics10030237.

70. Dobiasova H, Dolejska M. 2016. Prevalence and diversity of IncX plasmids carrying fluoroquinolone and β-lactam resistance genes in *Escherichia coli* originating from diverse sources and geographical areas. J Antimicrob Chemother 71:2118–2124. https://doi.org/10.1093/jac/dkw144.

71. Riley MA, Gordon DM. 1992. A survey of Col plasmids in natural isolates of *Escherichia coli* and an investigation into the stability of Col–plasmid Lineages. J Gen Microbiol 138:1345–1352. https://doi.org/10.1099/00221287-138-7-1345.

72. Sudhakari PA, Chandra B, Ramisetty M. 2021. Modeling colicin operons as genetic arms in plasmid-genome conflicts. bioRxiv 2021.05.15.444200. https://doi.org/10.1101/2021.05.15.444200.

73. Antipov D, Hartwick N, Shen M, Raiko M, Lapidus A, Pevzner PA. 2016. plasmidSPAdes: assembling plasmids from whole genome sequencing data. Bioinformatics 32:btw493. https://doi.org/10.1093/bioinformatics/btw493.

74. Arredondo-Alonso S, Willems RJ, van Schaik W, Schürch AC. 2017. On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data. Microb Genom 3. https://doi.org/10.1099/mgen.0.000128.

75. Lu M-C, Chen Y-T, Tang H-L, Liu Y-Y, Chen B-H, Wang Y-W, Chen Y-S, Teng R-H, Hong Y-P, Chiou C-S, Chiang M-K, Lai Y-C. 2020. Transmission and evolution of OXA-48-producing *Klebsiella pneumoniae* st11 in a single hospital in Taiwan. J Antimicrob Chemother 75:318–326. https://doi.org/10.1093/jac/dkz431.

76. Zhao S, Li C, Hsu C-H, Tyson GH, Strain E, Tate H, Tran T-T, Abbott J, McDermott PF. 2020. Comparative genomic analysis of 450 strains of salmonella enterica isolated from diseased animals. Genes (Basel) 11:1025. https://doi.org/10.3390/genes11091025.

77. Brenciani A, Morroni G, Pollini S, Tiberi E, Mingoia M, Varaldo PE, Rossolini GM, Giovanetti E. 2016. Characterization of novel conjugative multiresistance plasmids carrying *cfr* from linezolid-resistant *Staphylococcus epidermidis* clinical isolates from Italy. J Antimicrob Chemother 71:307–313. https://doi.org/10.1093/jac/dkv341.

78. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, Philippon A, Allesoe RL, Rebelo AR, Florensa AF, Fagelhauer L, Chakraborty T, Neumann B, Werner G, Bender JK, Stingl K, Nguyen M, Coppens J, Xavier BB, Malhotra-Kumar S, Westh H, Pinholt M, Anjum MF, Duggett NA, Kempf I, Nykäsenoja S, Olkkola S, Wieczorek K, Amaro A, Clemente L, Mossong J, Losch S, Ragimbeau C, Lund O, Aarestrup FM. 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother 75:3491–3500. https://doi.org/10.1093/jac/dkaa345.

79. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. J Clin Microbiol 52:1501–1510. https://doi.org/10.1128/JCM.03617-13.

80. Lawley TD, Klimke WA, Gubbins MJ, Frost LS. 2003. F factor conjugation is a true type IV secretion system. FEMS Microbiol Lett 224:1–15. https://doi.org/10.1016/S0378-1097(03)00430-0.

81. Christie PJ. 2016. the mosaic type IV secretion systems. EcoSal Plus 7. https://doi.org/10.1128/ecosalplus.ESP-0020-2015.

82. Morikawa K, Takemura AJ, Inose Y, Tsai M, Nguyen Thi LT, Ohta T, Msadek T. 2012. Expression of a cryptic secondary sigma factor gene unveils natural competence for DNA transformation in *Staphylococcus aureus*. PLoS Pathog 8:e1003003. https://doi.org/10.1371/journal.ppat.1003003.

83. Dunny GM. 2007. The peptide pheromone-inducible conjugation system of *Enterococcus faecalis* plasmid pCF10: cell–cell signalling, gene transfer, complexity and evolution. Philos Trans R Soc Lond B Biol Sci 362:1185–1193. https://doi.org/10.1098/rstb.2007.2043.

84. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

85. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 58:3895–3903. https://doi.org/10.1128/AAC.02412-14.

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

87. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

88. Prjibelski AD, Vasilinetc I, Bankevich A, Gurevich A, Krivosheeva T, Nurk S, Pham S, Korobeynikov A, Lapidus A, Pevzner PA. 2014. ExSPAnder: a universal repeat resolver for DNA fragment assembly. Bioinformatics 30:i293–i301. https://doi.org/10.1093/bioinformatics/btu266.

89. Geneious. 2021. Geneious v. 2021.0.3. https://www.geneious.com.

90. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.