Foodborne Origin and Local and Global Spread of *Staphylococcus saprophyticus* Causing Human Urinary Tract Infections

Opeyemi U. Lawal, Maria J. Fraqueza, Ons Bouchami, Peder Wøring, Mette D. Bartels, Maria L. Gonçalves, Paulo Paixão, Elsa Gonçalves, Cristina Toscano, Joanna Empel, Malgorzata Urbaś, M. Angeles Domínguez, Henrik Westh, Herminia de Lencastre, Maria Miragaia

*Staphylococcus saprophyticus* is a primary cause of community-acquired urinary tract infections (UTIs) in young women. *S. saprophyticus* colonizes humans and animals but basic features of its molecular epidemiology are undetermined. We conducted a phylogenomic analysis of 321 *S. saprophyticus* isolates collected from human UTIs worldwide during 1997–2017 and 232 isolates from human UTIs and the pig-processing chain in a confined region during 2016–2017. We found epidemiologic and genomic evidence that the meat-production chain is a major source of *S. saprophyticus* causing human UTIs; human microbiota is another possible origin. Pathogenic *S. saprophyticus* belonged to 2 lineages with distinctive genetic features that are globally and locally disseminated. Pangenome-wide approaches identified a strong association between pathogenicity and antimicrobial resistance, phages, platelet binding proteins, and an increased recombination rate. Our study provides insight into the origin, transmission, and population structure of pathogenic *S. saprophyticus* and identifies putative new virulence factors.

*Staphylococcus saprophyticus* is the cause of uncomplicated urinary tract infection (UTI) in 10%–20% of young women (1). Despite a greater successful treatment rate, *S. saprophyticus* UTI has a higher recurrent infection frequency than *Escherichia coli* UTI (2). Rare complications of *S. saprophyticus* UTI include acute pyelonephritis, nephrolithiasis, and endocarditis (3).

*S. saprophyticus* frequently colonizes humans and can be found in the gastrointestinal tract, vagina, and perineum (4). *S. saprophyticus* also is part of the gut and rectal flora of livestock, including pigs and cattle, and a frequent contaminant of meat and fermented food products (1). *S. saprophyticus* also has been recovered from polluted aquatic environments (5).

The reservoirs of *S. saprophyticus* causing UTI in humans are believed to be endogenous, but evidence is lacking. Moreover, given the frequent bacterial contamination of meat through the meat-processing chain, meat and food are speculated to be sources of human gut colonization and human *S. saprophyticus* infection (6). Some studies have reported high genetic diversity among isolates from human infections, food products, and other sources (5,7,8). However, previous studies were performed with a limited number of isolates, which prevented the description of the global and local molecular epidemiology of *S. saprophyticus*.

We used phenotypic, genomic, and pangenome-wide association study (pan-GWAS) approaches to characterize *S. saprophyticus* both globally and locally. In addition, we identified adaptive features that drive *S. saprophyticus* evolution, defined the *S. saprophyticus* population structure, investigated dissemination routes, and identified new pathogenicity factors.

Methods

Ethics Considerations

The human isolates used in our study were recovered as part of routine clinical diagnostic testing; thus,
ethics approval and informed consent were not required. All data were handled anonymously. Sample collection was in accordance with the European Parliament and Council decision for the epidemiologic surveillance and control of communicable disease through the European Antimicrobial Resistance Surveillance Network (http://www.ecdc.europa.eu/en/activities/surveillance/EARS-Net/Pages/index.aspx). Slaughterhouse samples were part of the routine control practices for evaluation of good hygiene practices and programs to assure meat safety (European Parliament and Council regulation no. 853/2004).

**Bacterial Isolates**

The global *S. saprophyticus* collection we used included 299 isolates from humans collected in 7 countries during 1997–2017: 286 from UTIs, 12 from invasive disease, and 1 from colonization (Appendix 1 Table 1, https://wwwnc.cdc.gov/EID/article/27/3/20-0852-App1.xlsx). We also analyzed the genomes of *S. saprophyticus* for 38 isolates from 5 other countries: 35 isolates from human UTIs (8), 2 from human hand swabs (8), an isolate from Byzantine Troy (8), and ATCC 15305 (9), a previously investigated human UTI-causing isolate.

The local collection included isolates collected in Lisbon, Portugal, during 2016–2017: 128 human UTI isolates collected in 3 hospitals and 104 slaughterhouse isolates collected from equipment, pork samples, workers’ hands, and a pig’s rectum. In addition, we included 5 isolates from animals and 12 isolates from food used in other studies (8) (Appendix 2, https://wwwnc.cdc.gov/EID/article/27/3/20-0852-App2.pdf).

**Whole-Genome Sequencing and Assembly**

**Phylogenetic Analysis**

We performed whole-genome sequencing (WGS) on MiSeq (Illumina, https://www.illumina.com) and MinIon nanopore (Oxford Nanopore, https://nanoporetech.com) platforms, as described (10) (Appendix 2). We separately analyzed global population and local epidemiology of *S. saprophyticus* and their phylogeny by using single-nucleotide polymorphisms (SNPs). We identified SNPs by mapping the draft genomes to a reference genome, *S. saprophyticus* ATCC 15305 (GenBank accession no. AP008934.1) by using the web-based CSI phylogeny version 1.4 (11) with the default parameter, but we disabled the minimum distance between SNPs in the parameter. We used Gubbins version 2.3.4 (12) with default parameters to concatenate SNPs and remove recombination regions. We reconstructed the phylogenies by using RAxML version 8.2.4 (https://github.com/stamatak/standard-RAXML) and generalized time-reversible nucleotide substitution with gamma correction model with 100 bootstrap value. We visualized the maximum-likelihood trees by using Interactive Tree of Life (https://itol.embl.de) (Figures 1–3; Appendix 2 Figure 3). Recombination to mutation (r/m) ratios detected by using Gubbins were calculated as the average r/m of isolates in the entire collection and separately for each lineage by using as reference closed genomes of KS40 for lineage G and KS160 for lineage S, both obtained on the MinIon platform (Oxford Nanopore).

**Pan-GWAS**

We used Prokka version 1.13 (https://vicbioinformatics.com/software.prokka.shtml) to annotate genomes and defined the pangenome by using 85% blastp (https://blast.ncbi.nlm.nih.gov/Blast.cgi) identity in Roary version 3.12 (http://sanger-pathogens.github.io/Roary). We performed GWAS by using Scoary version 1.6.16 (13) to identify genes associated with lineages and considered Bonferroni corrected p<0.05 (odds ratio [OR] >1) statistically significant; we identified genes associated with epidemiologic groups and considered Benjamini Hochberg corrected and pairwise comparison p<0.05 statistically significant. Sequence data from this study are available in GenBank (accession no. PRJNA604222).

**Resistome and Virulome Analyses**

We screened genomes for resistance and virulence genes by using ResFinder version 2.3 (http://cge.cbs.dtu.dk/services/ResFinder), an in-house virulence genes database, and the virulence factor database (14), integrated into ABRiccate version 0.5 (https://github.com/teemann/abricate). We considered genes with a threshold ≥90% nucleotide identity and ≥60% coverage to be present.

**Statistical Analyses**

We used Prism 6.0 (GraphPad, https://www.graphpad.com) to compare the means of 2 groups. We used a 2-tailed unpaired Mann-Whitney test or $\chi^2$ test for comparison and considered p<0.05 statistically significant.

**Results**

**Lineages of *S. saprophyticus* Causing UTIs**

Among study isolates, we first analyzed the diversity of *S. saprophyticus* causing UTIs by using genomic data of 321 human UTI isolates collected from 8 countries.
on 4 continents during 1997–2017. From the SNPs initially detected, 42% arose from recombination events in the population, corresponding to a mean genome-wide r/m of 1.5:1, meaning that high accumulation of SNPs was due to recombination rather than mutation in UTI strains. The maximum-likelihood tree constructed from the 9,134 SNPs without recombination defined 2 lineages, which we called G and S (Figure 1). Most (74%, 236/321) UTI isolates were from lineage G and differed by 0–4,318 SNPs with an average nucleotide identity (ANI) of 98.5%–99.999%, whereas S isolates (26%, n = 85/321) were slightly less distantly

Figure 1. Maximum-likelihood tree of Staphylococcus saprophyticus isolates recovered from human infections and colonization globally, 1997–2017. The tree was constructed by using 9,134 SNPs without recombination. Among analyzed isolates, 321 were recovered from UTIs, 12 from blood, and 4 from colonization. Each node represents a strain; nodes with identical color belong to the same lineage. The assembled contigs were mapped to the reference genome S. saprophyticus ATCC 15305 (GenBank accession no. AP008934.1; black star). Polymorphic sites resulting from recombination events in the single-nucleotide polymorphism (SNP) alignments were filtered out by using Gubbins version 2.3.4 (12). Maximum likelihood tree was reconstructed by using RAxML version 8.2.4 (https://github.com/stamatak/standard-RAXML). We performed generalized time-reversible nucleotide substitution model with gamma correction with 100 bootstraps random resampling for support. We visualized the tree by using Interactive Tree of Life (iTOL; https://itol.embl.de). Black triangles represent isolates fully sequenced by using the long-read nanopore technologies and used as reference to estimate r/m in the respective lineage. Cream color represents clusters G1, G2, G3, G4, and S1, which had dissemination and transmission in same country and in different countries. The outer ring represents isolates’ country of origin; blocks with identical color represent isolates from the same country. Of note, cluster G4 contains a pair of isolates collected in 2016 that had only 10 SNPs difference; one is a blood isolate from Barcelona, Spain (KS266) and the other is a UTI isolate recovered in Lisbon, Portugal (KS135). Scale bar indicates number of substitutions per site. UTI, urinary tract infection; r/m, recombination to mutation ratio.
Figure 2. Phylogenomic analysis and distribution of *Staphylococcus saprophyticus* isolates collected from human infections and a slaughterhouse, Portugal, 2016–2017. A) Maximum-likelihood tree of 232 isolates from human infections or slaughterhouse contamination. The tree was constructed by using 14,110 single-nucleotide polymorphisms (SNPs) without recombination. Each node represents a strain; nodes with identical color belong to the same lineage. The assembled contigs were mapped to the reference genome *S. saprophyticus* ATCC 15305 (GenBank accession no. AP008934.1; black star). SNPs generated from each genome were concatenated to single alignment corresponding to position of the reference genome. Polymorphic sites resulting from recombination events in the SNP alignments were filtered out by using Gubbins version 2.3.4 (12). Tree was reconstructed by using RAxML version 8.2.4 (https://github.com/stamatak/standard-RAxML). The generalized time-reversible nucleotide substitution model with gamma correction was performed with 100 bootstrap random re-samplings for support. The tree was visualized by using Interactive Tree of Life (iTOL; https://itol.embl.de). The clusters highlighted in cream represent admixture of isolates recovered from different sources that are closely related by SNPs in clusters G5–G11 and S2–S4. The inner ring (ring 1) represents genetic relatedness of isolates recovered from different sites inside the slaughterhouses and those recovered from infection in the community. The center ring (ring 2) identifies the isolates recovered from different hospitals. The outer rings (rings 3–8) represent the distribution of 6 genes that convey antimicrobial resistance. Scale bar indicates nucleotide substitutions per site. B) Source-based distribution of *S. saprophyticus* isolates in the lineage G and lineage S. Lineage G consisted isolates from infections, colonization, and contamination. Almost all (97%) isolates in lineage S are from human colonization and infection.
related, differing by 0–3,540 SNPs with an ANI of 99.3%–99.991% (Appendix 2 Figure 1).

*S. saprophyticus* lineages we identified among UTI isolates worldwide had distinctive features indicative of 2 evolutionary histories. Although strains of both lineages had similar genome size as determined by their closed genomes (lineage G was 2.5 Mb and S 2.6 Mb), isolates from lineage S (n = 6) grew significantly faster in tryptic soy broth incubated at 37°C (μ<sub>average</sub> = 0.34 h<sup>-1</sup>) than G strains (n = 8) (μ<sub>average</sub> = 0.22 h<sup>-1</sup>; p = 0.0007) (Figure 4; Appendix 2). Moreover, we separately analyzed r/m of both lineages by using the respective closed genome of strains from each lineage as a reference and found that the mean estimated r/m was 9 times higher in S isolates (r/m = 4.4:1) than G isolates (r/m = 0.5:1). We did not detect a temporal signal when performing regression analysis (r) of tip-to-root distance versus isolation date, either in the entire collection (r = −0.2423) or for separate lineages (for G, r = −0.1314; for S, r = 0.1889) (Appendix). Hence, we could not determine when the 2 lineages diverged. The lack of temporal signal probably results from the limited number of isolates in each time point.

To further compare the 2 lineages, we constructed the pangenome of the 338 human *S. saprophyticus* genomes and identified 10,222 genes with 85% blastp clustering by using Roary. Among these, 8,351 genes,
present in <99% of the genomes, constituted the accessory genome. A gene accumulation plot of all the genes against the genomes sequenced showed that *S. saprophyticus* has an open pangenome (Appendix 2 Figure 2, panels A, B). The pan-GWAS analysis of the accessory genome indicated that G and S isolates have different genomic content (Bonferroni p<0.05, OR >1). A total of 128 genes were specific/enriched in the G lineage, including those encoding a type I restriction subunit (*group_383*), a defense mechanism against genetic transfer (15); metabolism of melibiose (*ebgA*, *melB*) (16) and inositol (*iolE*) (17), compounds that are excreted in urine; toxin–antitoxin systems (*group_4685*, *group_5665*), involved in stress response (18); and antimicrobial resistance (*qacA*) (19) (Table 1; Figure 5, panel A; Appendix 2 Table 1).

For S lineage, 237 genes were specific/enriched, some of which are involved in metabolism...
of ascorbate (*ulaA*) and thiamine (*tenl*) (20); induction of phosphate starvation (*psie*) , which was previously linked to switch on virulence in other uropathogens (21); platelet binding (*splE, sdeE*) associated with virulence (22); steroid metabolism (*group_7190*); and resistance to trimethoprim (*dfrG*) and biocides (*qacC*) (Table 2; Figure 5, panel A; Appendix 2 Table 1). Additional genes associated with S lineage that could explain its increased growth rate include those linked to lactose and cellobiose metabolism (*group_5572* gene) (23) and cell wall hydrolysis (*lytN*) (24), whereas the prevalence of transposases (*group_3547* and *group_1828*) (25) and phage genes (*recT* and *yueB*) (27) could justify its increased recombination rate.

### Local and Global Spread of *S. saprophyticus*

#### Clones Causing UTIs

We observed no time-based clustering of *S. saprophyticus* UTI isolates in the phylogeny, but we noted some geographic clustering. In particular, 89% (n = 78) of S isolates were found in Portugal and Spain (Figure 1). In addition, we identified clusters containing isolates from a single country. For instance, cluster G3 (1–8 SNPs) contained only isolates from Portugal and cluster G1 (0–63 SNPs) only isolates from Denmark (Figure 1).

We noted a high degree of isolate admixture in the maximum-likelihood tree of our global collection, suggesting that *S. saprophyticus* isolates of both lineages are disseminated geographically. G strains were distributed most widely, in 11 countries on 4 continents, whereas we found S isolates in only 6 countries. Despite the genetic diversity described, we still found isolates from different countries that differed by only a few SNPs. One pair in cluster G4 that had only 10 SNPs difference was a blood isolate from Barcelona, Spain (KS266), and a UTI isolate from Lisbon (KS135), collected in 2016 (Figure 1).

Although a relatedness cutoff is not defined yet for *S. saprophyticus*, the low number of SNPs observed between strains from the same and different countries is below the relatedness cutoff of 10–40 SNPs for most bacterial species (28). The apparent relatedness we noted implies that UTI isolates from different patients in the same country and in different countries are highly related and could belong to a cross-border chain of transmission. This finding challenges the assumption that *S. saprophyticus* causing UTIs were mainly endogenous (29).

#### Genetic Relatedness of *S. saprophyticus* from Slaughterhouses and UTIs

Pork is the most frequently consumed red meat in Europe (30) and is often contaminated with *S. saprophyticus* (1). We found that 35% of slaughterhouse samples (from meat, equipment, workers’ hands, and a live pig) were contaminated with *S. saprophyticus*.

To understand whether *S. saprophyticus* causing UTIs could be related to *S. saprophyticus* in pork, we compared 104 isolates collected from a slaughterhouse against 128 isolates collected from human UTIs in Lisbon during 2016–2017. Among the 104 isolates from the slaughterhouse, 39 (37.5%) were collected from meat, 32 (30.8%) from equipment, 32 (30.8%) from workers’ hands, and 1 (1%) from a live pig. SNP-based phylogenetic analysis with a tree-rooted at the midpoint showed that most (91%; 95/104) slaughterhouse isolates belonged to lineage G (Figure 2, panels A, B) and that a strain from slaughterhouse equipment was at the base of this lineage (bootstrap 100). In addition, the phylogenetic reconstruction including isolates from this study and other isolates

### Table 1. List of genes exclusively associated with lineage G *Staphylococcus saprophyticus* strains in study of foodborne origin and local and global spread of *S. saprophyticus* causing human urinary tract infections*

<table>
<thead>
<tr>
<th>Genes</th>
<th>Gene predicted function</th>
<th>Biologic function group</th>
<th>Frequency, %</th>
<th>Reference no.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>group_383</em></td>
<td>Type I site-specific deoxyribonuclease</td>
<td>Endonuclease</td>
<td>96</td>
<td>(15)</td>
</tr>
<tr>
<td><em>icaR</em></td>
<td>Ica operon HTH-type negative transcription regulator</td>
<td></td>
<td>95</td>
<td>NA</td>
</tr>
<tr>
<td><em>licT</em></td>
<td>Transcription antiterminator LicT</td>
<td>Transcription</td>
<td>80</td>
<td>NA</td>
</tr>
<tr>
<td><em>iolE</em></td>
<td>Inosine dehydratase</td>
<td>Inositol metabolism</td>
<td>86</td>
<td>(17)</td>
</tr>
<tr>
<td><em>group_4976</em></td>
<td>Phage infection protein</td>
<td>Phage-related protein</td>
<td>46</td>
<td>NA</td>
</tr>
<tr>
<td><em>group_869</em></td>
<td>Bacteriophage integrase</td>
<td>Phage-related protein</td>
<td>42</td>
<td>NA</td>
</tr>
<tr>
<td><em>tnpC_1</em></td>
<td>Transposase for transposon Tn554</td>
<td>Mobile genetic element</td>
<td>17</td>
<td>NA</td>
</tr>
<tr>
<td><em>group_744</em></td>
<td>Putative glucarate transporter</td>
<td>Transporter</td>
<td>29</td>
<td>NA</td>
</tr>
<tr>
<td><em>group_4685</em></td>
<td>Addiction module toxin Txe/YoeB family protein</td>
<td>Stress response (Type II toxin–antitoxin system)</td>
<td>20</td>
<td>(18)</td>
</tr>
<tr>
<td><em>group_5665</em></td>
<td>Addiction module antitoxin Axe family protein</td>
<td>Stress response (Type II toxin–antitoxin system)</td>
<td>19</td>
<td>(18)</td>
</tr>
<tr>
<td><em>yhhQ</em></td>
<td>Putative cysteine-rich protein YhhQ</td>
<td>Putative function</td>
<td>17</td>
<td>NA</td>
</tr>
</tbody>
</table>

*List of genes also include 53 hypothetical proteins. Bonferroni adjusted p<0.032. NA, not applicable.
from production and companion animals (including 2 pigs, 2 bovine, and 1 canine) and food (8) showed that most (3/5) animal isolates clustered together at a basal clad of lineage G (bootstrap 100) (Appendix 2 Figure 2). Some clusters in the phylogenetic tree (e.g., G9) had slaughterhouse isolates at the base and UTI isolates at the tip. However, we also observed the opposite (e.g., G11), tree clusters with slaughterhouse isolates at the tip and UTI isolates at the base (Figure 2, panel A). Moreover, 41% of G strains included the antimicrobial resistance gene tetK (p<0.0001) (Figure 2, panel A), which is associated with resistance to tetracycline, an antimicrobial drug extensively used in animal production (31).

Phylogenetic reconstruction of all isolates from Lisbon based on SNPs provided additional examples of admixture of isolates recovered from different sampling sites in the slaughterhouse and from the slaughterhouse and humans. Isolates from meat were frequently intermixed with isolates from equipment and colonized workers as observed in cluster G6 wherein strains differed by only 1–65 SNPs (Figure 2, panel A; Figure 3). Likewise, cluster G9 included isolates from the slaughterhouse that were intermixed with human UTI isolates. Isolates collected from slaughterhouse equipment differed by only 19 SNPs from human UTI isolates. Likewise, human UTI isolates differed from meat isolates by 25 SNPs and from isolates of colonized workers by 26 SNPs. In addition, slaughterhouse and human UTI isolates had the same antimicrobial resistance profile, exhibiting resistance to fosfomycin, fusidic acid, and tetracycline (Appendix

---

**Figure 5.** Genetic determinants that contribute to the distinction of clonal lineages and lifestyle of *Staphylococcus saprophyticus*. The graph displays determinants that contribute (A) and mediate (B) adaptation of *S. saprophyticus* to either infection or contamination. We used the genome-wide association study (GWAS) method to identify genetic factors by using 2 association comparisons: lineage G versus lineage S and human infection versus surface contamination. We used the pairwise comparison and included a core-SNP phylogenetic tree without recombination to remove the lineage effect in the analysis. Hits with Benjamini Hochberg corrected p < 0.05 and odds ratio >1 were considered statistically significant. We grouped the identified genes into biologic functions based on gene annotation predicted by Prokka (https://vicbioinformatics.com/software.prokka.shtml). Some genetic factors that were associated with infections and contamination also were associated with the lineages despite subjecting the GWAS to lineage correction.
In contrast to isolates belonging to lineage G, which were mostly from UTIs and the slaughterhouse, S isolates were almost exclusively of human origin (97%; n = 66/68), either from UTIs (n = 59) or human colonization (n = 7) (Figure 2, panels A, B). When we reconstructed the phylogeny of all isolates in this and other studies (8) (Appendix 2 Figure 3), a human isolate was at the base of the S lineage. The only 2 S isolates seen in animals were from nonhuman primates. Furthermore, a resistance determinant for trimethoprin (drfG), which routinely is used to treat human UTIs, was associated with this lineage (18%; p < 0.05) (Figure 2, panel A; Table 2).

To determine whether S isolates could have originated in humans, we grew isolates from both lineages in the absence and presence of human physiologic concentrations (15–350 pg/mL) of estradiol (32), a female hormone commonly found in urine and the vagina, and in different pH values mimicking the stomach (pH 2.5), vagina (pH 4.5) (33), skin (pH 5.5), and urine (pH 4.5–8.0) (34). The growth rate of S isolates did not change significantly at the highest physiologic concentration of estradiol specific to humans (0.34 h⁻¹ vs. 0.33 h⁻¹), but the growth rate for G strains decreased by 59% at this concentration (0.22 h⁻¹ vs. 0.13 h⁻¹; p = 0.0007) (Figure 4, panels A–C; Appendix 2). All isolates grew at all pH levels assayed, except for pH 2.5. At pH 4.5 and 5.5, isolates of both lineages had similar growth rates, but S isolates had a higher growth rate than G isolates at pH 8, although this difference was not statistically significant (p = 0.133). These results suggest that lineage S isolates are more adapted than lineage G strains to high estradiol concentrations found in women, but not found in other female animal hosts, such as pigs, bovine, or canines (35). GWAS also identified a gene involved in

### Table 2. List of lineage S genes exclusively associated with *Staphylococcus saprophyticus* strains in study of foodborne origin and local and global spread of *S. saprophyticus* causing human urinary tract infections*

<table>
<thead>
<tr>
<th>Genes</th>
<th>Gene predicted function</th>
<th>Biologic function group</th>
<th>Frequency, %</th>
<th>Reference no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>group_5572</td>
<td>Phosphotransferase system, lactose/cellulose-specific IIC component</td>
<td>Sugar metabolism</td>
<td>99</td>
<td>(23)</td>
</tr>
<tr>
<td>ullaA</td>
<td>PTS system ascorbate-specific IIC component</td>
<td>Sugar metabolism</td>
<td>99</td>
<td>(20)</td>
</tr>
<tr>
<td>group_5955</td>
<td>Sugar phosphate isomerase/epimerase</td>
<td>Sugar metabolism</td>
<td>26</td>
<td>(23)</td>
</tr>
<tr>
<td>lytN</td>
<td>C51 family D-Ala-D-Gly carboxypeptidase</td>
<td>Cell wall hydrolase</td>
<td>85</td>
<td>(24)</td>
</tr>
<tr>
<td>psIE</td>
<td>Protein PsIE</td>
<td>Phosphate starvation</td>
<td>22</td>
<td>(21)</td>
</tr>
<tr>
<td>group_1438</td>
<td>Arsinite methyltransferase</td>
<td>Arsinite resistance</td>
<td>25</td>
<td>NA</td>
</tr>
<tr>
<td>dfrG</td>
<td>Tricentreprin-resistance dihydrofolate reductase</td>
<td>Antimicrobial resistance</td>
<td>18</td>
<td>NA</td>
</tr>
<tr>
<td>group_7190</td>
<td>3-hydroxyxystroid dehydrogenase/isomerase</td>
<td>Steroid metabolism</td>
<td>17</td>
<td>(26)</td>
</tr>
<tr>
<td>group_273</td>
<td>Recombinase/resolvase</td>
<td>Mobile genetic element</td>
<td>69</td>
<td>(25)</td>
</tr>
<tr>
<td>group_2198</td>
<td>Putative ATPase/transposase</td>
<td>Mobile genetic element</td>
<td>26</td>
<td>(25)</td>
</tr>
<tr>
<td>group_275</td>
<td>Recombinase/resolvase</td>
<td>Mobile genetic element</td>
<td>21</td>
<td>(25)</td>
</tr>
<tr>
<td>group_355</td>
<td>Transposase for IS431mec</td>
<td>Mobile genetic element</td>
<td>17</td>
<td>(25)</td>
</tr>
<tr>
<td>group_7470</td>
<td>Putative replication-associated protein</td>
<td>Mobile genetic element</td>
<td>10</td>
<td>NA</td>
</tr>
<tr>
<td>group_3363</td>
<td>Putative DoxX family membrane protein</td>
<td>Putative functions</td>
<td>85</td>
<td>NA</td>
</tr>
<tr>
<td>mvIM</td>
<td>NADH-dependent dehydrogenase</td>
<td>Putative functions</td>
<td>26</td>
<td>NA</td>
</tr>
<tr>
<td>nmRA</td>
<td>Putative nmRA negative transcription</td>
<td>Transcription</td>
<td>17</td>
<td>NA</td>
</tr>
<tr>
<td>group_7195</td>
<td>Amidohydrolase</td>
<td>Hydrolase</td>
<td>17</td>
<td>NA</td>
</tr>
<tr>
<td>group_6430</td>
<td>Putative restriction enzyme</td>
<td>Restriction enzyme</td>
<td>17</td>
<td>NA</td>
</tr>
</tbody>
</table>

*List of genes also include 84 hypothetical proteins. Bonferroni p < 0.031. NA, not applicable.*
steroid metabolism, 3-β hydroxysteroid dehydroge-
nase (HSD), associated with lineage S (Table 2; Figure 5, panels A, B). Steroids such as estradiol are primary
signaling molecules for host–microbe interactions (26) and involved in the interconversion of active
and inactive steroid hormones (26). Presence of HSD
could be an adaptive evolution to colonization of
the bladder, a hormone-rich environment. Evidence
supports a human (primate) origin for lineage S, but
studies sampling a wider range of ecologic sites and
geographic regions are needed.

UTIs and S. saprophyticus Dissemination
among Humans in the Community
To explore dissemination of S. saprophyticus causing
UTIs in the community through human-to-human
contact, we analyzed genomic data of the 128 UTI
isolates from outpatients of 3 hospitals in the Lis-
bon area. Transmission of S. saprophyticus from lin-
eage G and S between persons in the community was
apparent, as demonstrated by the high relatedness
of strains from UTI patients at different hospitals.
In the G cluster, G10 isolates had 9–24 SNPs differ-
ence, and in the S cluster, S2, differed by 6–64 SNPs
(Figure 2, panel A). However, due to data protection
regulations, we could not ascertain whether the pa-
tients were epidemiologically linked. Results suggest
that patients might have acquired these strains from
the same reservoir or that direct and indirect cross-
transmission could have occurred in the community.

Disease Signatures among S. saprophyticus Populations
Several virulence factors, including urease, have
been described in S. saprophyticus (36), but the basis
of pathogenicity in this species is mainly unknown.
We used a pan-GWAS approach to compare the ge-
etic content of 128 isolates from human infections
to 104 isolates recovered from a slaughterhouse, all
collected in Lisbon during 2016–2017. We identi-
fied 6 genes that appear to be associated with an in-
creased pathogenic potential in S. saprophyticus (Fig-
ure 5, panel A; Appendix 2 Tables 2–5). These genes
included those encoding an SplE-like protein and a
gene cluster encoding a complete accessory secre-
tory system associated with a serine-rich adhesion,
similar to SraP. Previous studies have described
highly homologous secretory system (>97% nucleo-
tide identity) associated to serine-rich proteins that
bind platelets, including SraP in S. aureus (37) and
UafB in S. saprophyticus (38).

Figure 6. A proposed model for the dissemination and transmission of Staphylococcus saprophyticus in the community. The arrows
represent the dissemination and transmission of S. saprophyticus isolates that belonged to lineage G (green) and lineage S (yellow).
Lineage G S. saprophyticus strains are of animal origin and enter the slaughterhouse through production animals, such as pigs, persist
on the equipment, and contaminate the meat in processing chain. Lineage G strains could enter the community through contaminated
meat and workers colonized in the slaughterhouse. Lineage S strains most likely are of humans and primate origin and probably are
disseminated by person-to-person contact within the community.
We also found other genes associated with infection that encoded phage proteins. An analysis on 128 UTI isolates using PHASTER (39) identified 5 phages in 48 (38%) strains. The phages were 38–125 Kb and had <50% identity with any known phage. The genomic vicinity of the identified phages varied in the chromosome, suggesting that lysogenic conversion was not the mechanism involved in pathogenicity of these strains; instead, hypothetical phage genes could be crucial for pathogenicity. In addition, genes encoding resistance to the antimicrobial drugs trimethoprim (dfrG), lincosamide (InhA), streptogramin B (erm44v), or to macrolides (mphC-msrA) also were associated with strains from human infections (Figure 2; Figure 5, panel B; Appendix 2 Tables 2–5).

Another factor highly associated with human infections was the occurrence of recombination, which was 6 times higher in isolates from UTIs (r/m 1.7:1) than from isolates associated with colonization or contamination (r/m 0.3:1). This finding suggests recombination might be a strategy of *S. saprophyticus* to evade the host immune system, as described elsewhere (40).

**Discussion**

We identified 2 *S. saprophyticus* lineages, G and S, associated with human UTIs that appeared to have different evolutionary histories. Our data support a foodborne origin for lineage G and its transmission through food products to humans. We also found evidence of a human origin for lineage S and additional proof for the occurrence of direct or indirect human-to-human dissemination in the community, which could explain not only the local dissemination of both lineages but also the wide geographic dissemination, as described for *S. aureus* (41) and *E. coli* (42).

Our conclusions were limited by the characteristics of the sample collection analyzed. In particular, the lack of temporal signal did not enable inference of the direction of transmission between the sampling sites. We could establish obvious genomic relatedness between the meat production chain and human UTIs and UTIs from different persons in the community. However, lack of samples from other animal hosts, environments, or different ecologic niches in humans, did not enable us to establish pig meat and humans as the unique sources of the 2 lineages nor identify the preferred ecologic niche of *S. saprophyticus* in humans.

Our observed *S. saprophyticus* colonization rate among pigs was extremely low (1%) compared with previous studies (43), which could be explained by possible cross-inhibitory bacterial interactions (44); however, contamination within the slaughterhouse environment sometimes reached 35%. Amplification of *S. saprophyticus* in the slaughterhouse environment was probably potentiated by resistance to biocides. We found that slaughterhouse workers’ washed hands were colonized with strains that were highly related to slaughterhouse environmental isolates and to strains causing UTIs, suggesting exposure of workers to the slaughterhouse environment as a risk factor for human colonization (Figure 6).

The transmission chain of S lineage isolates appears to be different and to have no link to meat production. S lineage UTIs might originate from human gut or vaginal colonization; both have been reported as possible human niches (2,4). Evidence for the human origin of S lineage included the almost exclusive (97%) identification in humans, where S lineage was better adapted to the physiologic high concentration of human female sex hormones; grew at vagina, skin, and urine pH values; and harbored 2 serine-proteases that presumably can bind to human platelets, as described for SraP (37). The distinct genetic content and phenotypic features of lineages G and S further reflected the diverse selective pressures of the human and animal, slaughterhouse, and food environments and suggest different evolutionary strategies toward pathogenicity. In particular, we found that tetracycline and antimicrobial resistance genes associated with colonization-contamination also were associated with isolates of the G lineage. Furthermore, genes encoding sugar metabolism, serine proteases, and a trimethoprim resistance associated with infection were also associated with isolates of the S lineage, implying distinctive specialization of the 2 lineages.

Our results also indicate a key role of setting-associated antimicrobial drug usage, especially for trimethoprim, macrolides, and tetracycline, in resistance development and pathogenicity. Subinhibitory concentrations of these drugs have been shown to promote virulence in bacteria through the induction of biofilm production (45), quorum-sensing (45), or phages (46) and might also increase *S. saprophyticus* pathogenicity.

We found a high r/m rate in isolates of the S lineage and in strains from UTIs, comparable to naturally transformable bacterial species like *Klebsiella pneumoniae* and *Streptococcus pyogenes* (47). A similar phenomenon was observed for *S. epidermidis* (48); variation in cell surface proteins was shown to contribute to evasion of human immunity (48), and a similar strategy might be advantageous for *S. saprophyticus* in infection. The recombination might have resulted from defects in repair of double stranded DNA breaks from...
oxidative stress induced by leukocytes during infection, as previously described (48). However, the mechanism linking infection and recombination in S. saprophyticus remains elusive.

Last, we identified factors associated with infection that could represent new S. saprophyticus virulence factors, including 2 serine-proteases, sraP-like and splE-like, and phages SraP and SplE, which previously were connected to pathogenicity in S. aureus (22,37). In addition, phages have been described to transport pathogenicity islands in staphylococci (49). Our study constitutes a deep-structured analysis of S. saprophyticus population structure and genomic epidemiology, providing groundwork for future studies on the pathogenicity and population genetics of this bacterium.

O.U.L. was supported by PhD grants from the Fundação para a Ciência e Tecnologia (FCT) (grant no. PD/BD/113992/2015). This work was partially supported by FCT (project no. PTDC/CVT-CVT/29510/2017), Microbiologia Molecular, Estrutural e Celular (project nos. LISBOA-01-0145-FEDER-007660 and UID/Multi/04378/2019) and by the COMPETE2020 Programa Operacional Competitividade e Internacionalização; ONEIDA (project no. LISBOA-01-0145-FEDER-016417), and co-funded by Fundos Europeus Estruturais e de Investimento from Programa Operacional Regional Lisboa2020 and by national funds through FCT.


About the Author

Dr. Lawal is a postdoctoral researcher at the Instituto de Tecnologia Quimica e Biologica, Universidade Nova de Lisboa (ITQB-NOVA), Oeiras, Portugal. His primary research interests include antimicrobial resistance, bacterial evolution and transmission dynamics of Staphylococcus saprophyticus.

References


17. Dinicola S, Minini M, Unter V, Verna R, Cucina A, Bizzarri M. Nutritional and acquired deficiencies in inositol bioavail-

j.1574-6968.2011.02218.x


20. Costilow ZA, Degnan PH. Thiamine acquisition strategies impact metabolism and competition in the gut microbe

21. Chekabab SM, Harel J, Dozois CM. Interplay between genetic regulation of phosphate homeostasis and bacterial viru-

1371/journal.pone.0038407

23. Ghali I, Sofyan A, Ohmori H, Shinkai T, Mitsumori M. Diauxic growth of Fibrobacter succinogenes S55 on cellobiose

24. Franken MB, Hendricks APA, Missiakas DM, Schneewind O. LytN, a murein hydrolase in the cross-wall compartment of
https://doi.org/10.1074/jbc.M111.258863

th00349.x

26. García-Gómez E, González-Pedrajo B, Camacho-Arroyo I. Regulation of phosphate homeostasis and bacterial viru-

27. Datta S, Costantino N, Zhou X, Court DL. Identification and analysis of recombineering functions from Gram-negative
0709089105

28. Schürch AC, Arredondo-Alonso S, Willems RJL, Goering RV. Whole genome sequencing options for bacterial strain

https://doi.org/10.1007/mda.2003.8

https://doi.org/10.1016/j.jclepro.2015.03.048


https://doi.org/10.1111/j.1574-3164.2010.00896.x


38. King NP, Beatson SA, Totsika M, Ulett GC, Alm RA, Manning PA, et al. UafB is a serine-rich repeat adhesin of
Staphylococcus saprophyticus that mediates binding to fibronectin, fibrinogen and human uroepithelial cells. Microbiology (Reading). 2011;157:1161–75. https://doi.org/10.1099/mic.0.04769-0

39. Arndt D, Grant JR, Marcus A, Sajed T, Alm RA, et al. PHASTER: a better, faster version of the PHAST phage

40. Yahara K, Didelot X, Jolley KA, Kobayashi I, Maiden MCJ, et al. PHASTER: a better, faster version of the PHAST phage


42. Vincent C, Boerlin P, Daignault D, Dozois CM, Dutil L, et al. Phage typing and epidemiologic analysis based on single

43. Arndt D, Grant JR, Marcus A, Sajed T, Alm RA, et al. PHASTER: a better, faster version of the PHAST phage

44. Verstappen KM, Willems E, Fluit AC, Duim B, Martens M, Wagenaar JA. Staphylococcus aureus nasal colonization

Foodborne Origin of S. saprophyticus


Address for correspondence: Maria Miragaia, Laboratory of Bacterial Evolution and Molecular Epidemiology, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa (ITQB-NOVA), Avenida da República, 2780-157, Oeiras, Portugal; email: miragaia@itqb.unl.pt

April 2020

Vectorborne Infections

- Stemming the Rising Tide of Human-Biting Ticks and Tickborne Diseases, United States
- Imported Arbovirus Infections in Spain, 2009–2018
- Decreased Susceptibility to Azithromycin in Clinical Shigella Isolates Associated with HIV and Sexually Transmitted Bacterial Diseases, Minnesota, USA, 2012–2015
- High Incidence of Active Tuberculosis in Asylum Seekers from Eritrea and Somalia in the First 5 Years after Arrival in the Netherlands
- Severe Dengue Epidemic, Sri Lanka, 2017
- Severe Fever with Thrombocytopenia Syndrome, Japan, 2013–2017
- Comprehensive Profiling of Zika Virus Risk with Natural and Artificial Mitigating Strategies, United States
- Genomic Insight into the Spread of Meropenem-Resistant Streptococcus pneumoniae Spain-ST81, Taiwan
- Isolation of Drug-Resistant Gallibacterium anatis from Calves with Unresponsive Bronchopneumonia, Belgium
- Guaroa Virus and Plasmidium vivax Co-Infections, Peruvian Amazon
- Intensified Short Symptom Screening Program for Dengue Infection during Pregnancy, India
- Prevalence of Antibodies to Crimean-Congo Hemorrhagic Fever Virus in Ruminants, Nigeria, 2015
- Recurrent Herpes Simplex Virus 2 Lymphocytic Meningitis in Patient with IgG Subclass 2 Deficiency
- Health-Related Quality of Life after Dengue Fever, Morelos, Mexico, 2016–2017
- Person-to-Person Transmission of Andes Virus in Hantavirus Pulmonary Syndrome, Argentina, 2014
- Ebola Virus Neutralizing Antibodies in Dogs from Sierra Leone, 2017
- Outbreak of Dirkmeia churashimaensis Fungemia in a Neonatal Intensive Care Unit, India
- Rift Valley Fever Outbreak, Mayotte, France, 2018–2019
- Detection of Zoonotic Bartonella Pathogens in Rabbit Fleas, Colorado, USA
- Human-to-Human Transmission of Monkeypox Virus, United Kingdom, October 2018
- Whole-Genome Analysis of Salmonella enterica Serovar Enteritidis Isolates in Outbreak Linked to Online Food Delivery, Shenzhen, China, 2018
- Pruritic Cutaneous Nematodiasis Caused by Avian Eyeworm Oxyspirura Larvae, Vietnam
- Novel Rapid Test for Detecting Carbapenemase
- Arthritis Caused by MRSA CC398 in a Patient without Animal Contact, Japan
- Detection of Rocio Virus SPH 34675 during Dengue Epidemics, Brazil, 2011–2013
- Epidemiology of Lassa Fever and Factors Associated with Deaths, Bauchi State, Nigeria, 2015–2018
- Plague Epicentriic Dynamics in Chipmunk Fleas, Sierra Nevada Mountains, California, USA, 2013–2015
- Knowledge of Infectious Disease Specialists Regarding Aspergillosis Complicating Influenza, United States

To revisit the April 2020 issue, go to: https://wwwnc.cdc.gov/eid/articles/issue/26/4/table-of-contents