



Staphylococcal Protein A (*spa*) Locus Is a Hot Spot for Recombination and Horizontal Gene Transfer in *Staphylococcus pseudintermedius*

Alem Zukancic,^a Mubin A. Khan,^b Sumayya J. Gurmen,^a Quinn M. Gliniecki,^a Dayna L. Moritz-Kinkade,^a Carol W. Maddox,^a Md Tauqeer Alam^{a,c,d}

^aDepartment of Pathobiology, College of Veterinary Medicine, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

^bSchool of Information Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

^cDepartment of Veterinary Clinical Medicine, College of Veterinary Medicine, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

^dInfection Genomics for One Health Theme, Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

ABSTRACT *Staphylococcus pseudintermedius* is a major canine pathogen but also occasionally colonizes and infects humans. Multidrug-resistant methicillin-resistant *S. pseudintermedius* (MDR MRSP) strains have emerged globally, making treatment and control of this pathogen challenging. Sequence type 71 (ST71), ST68, and ST45 are the most widespread and successful MDR MRSP clones. The potential genetic factors underlying the clonal success of these and other predominant clones remain unknown. Characterization of the pangenome, lineage-associated accessory genes, and genes acquired through horizontal gene transfer from other bacteria is important for identifying such factors. Here, we analyzed genome sequence data from 622 *S. pseudintermedius* isolates to investigate the evolution of pathogenicity across lineages. We show that the predominant clones carry one or more lineage-associated virulence genes. The gene encoding staphylococcal protein A (SpA), a key virulence factor involved in immune evasion and a potential vaccine antigen, is deleted in 62% of isolates. Most importantly, we have discovered that the *spa* locus is a hot spot for recombination and horizontal gene transfer in *S. pseudintermedius*, where genes related to restriction modification, prophage immunity, mercury resistance, and nucleotide and carbohydrate metabolism have been acquired in different lineages. Our study also establishes that ST45 is composed of two distinct sublineages that differ in their accessory gene content and virulence potential. Collectively, this study reports several previously undetected lineage-associated genetic factors that may have a role in the clonal success of the major MDR MRSP clones. These data provide a framework for future experimental studies on *S. pseudintermedius* pathogenesis and for developing novel therapeutics against this pathogen.

IMPORTANCE *Staphylococcus pseudintermedius* is a major canine pathogen but can also occasionally infect humans. Identification of genetic factors contributing to the virulence and clonal success of multidrug-resistant *S. pseudintermedius* clones is critical for the development of therapeutics against this pathogen. Here, we characterized the genome sequences of a global collection of 622 *S. pseudintermedius* isolates. We show that all major clones, besides carrying core virulence genes, which are present in all strains, carry one or more lineage-specific genes. Many of these genes have been acquired from other bacterial species through a horizontal gene transfer mechanism. Importantly, we have discovered that the staphylococcal protein A gene (*spa*), a widely used marker for molecular typing of *S. pseudintermedius* strains and a potential vaccine candidate antigen, is deleted in 62% of strains. Furthermore, the *spa* locus in *S. pseudintermedius* acts as a reservoir to accumulate lineage-associated genes with adaptive functions.

Citation Zukancic A, Khan MA, Gurmen SJ, Gliniecki QM, Moritz-Kinkade DL, Maddox CW, Alam MT. 2020. Staphylococcal protein A (*spa*) locus is a hot spot for recombination and horizontal gene transfer in *Staphylococcus pseudintermedius*. mSphere 5:e00666-20. <https://doi.org/10.1128/mSphere.00666-20>.

Editor Paul D. Fey, University of Nebraska Medical Center

Copyright © 2020 Zukancic et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Md Tauqeer Alam, mtalam@illinois.edu.

Received 1 July 2020

Accepted 8 October 2020

Published 28 October 2020

KEYWORDS *Staphylococcus*, antibiotic resistance, genomics, horizontal gene transfer

Staphylococcus pseudintermedius is a leading cause of pyoderma, otitis externa, urinary tract infections, and postsurgical abscesses in dogs. It also occasionally infects humans, especially the pet owners and veterinarians in close contact with the infected animals. Almost all human cases of *S. pseudintermedius* infections are reported to be zoonotic transmissions from the infected companion animals, indicating that the same strain can infect both animals and humans (1–4). In humans, it has been associated with skin infections, ulcers, brain abscess, cardioverter defibrillator infection, bacteremia, cellulitis, endocarditis, and chronic rhinosinusitis (5–8). The emergence and global spread of multidrug-resistant methicillin-resistant *S. pseudintermedius* (MDR MRSP) strains poses a serious challenge for the treatment and control of this important pathogen (9–11). The pattern of the emergence and evolution of antibiotic resistance in *S. pseudintermedius* is very similar to the pattern observed in *Staphylococcus aureus*, with both species acquiring almost identical types of resistance against commonly used β -lactam and non- β -lactam antibiotics, including fluoroquinolones (9, 12). Zoonotic transmission of MDR MRSP strains also poses a high potential risk of horizontal transfer of antibiotic resistance genes to other bacterial species within the human host. It is worth noting that the *mecA* gene, carried by the staphylococcal cassette chromosome *mec* (SCC*mec*) element in methicillin-resistant *S. aureus* (MRSA), is thought to have been acquired from an animal-related commensal *Staphylococcus* species, *Staphylococcus fleurettii*, through horizontal gene transfer (HGT) (13). Horizontal transfer of the *mecA* locus from *S. fleurettii* converted methicillin-susceptible *S. aureus* (MSSA) into MRSA (13).

The population structure of *S. pseudintermedius* is heterogenous, with over 2,000 sequence types (STs) recorded to date in the PubMLST database (<https://pubmlst.org/organisms/staphylococcus-pseudintermedius/>). Some of these STs, such as ST71, ST68, and ST45, the predominant clones of Europe, North America, and Asia, respectively, have spread globally, becoming the most widespread MDR MRSP clones (9, 14). The emergence and clonal expansion of MDR MRSP clones ST496, ST181, and ST258 have also been observed in some countries (15, 16). The enhanced fitness and clonal success of the epidemic clones is generally associated with the acquisition of antibiotic resistance, mobile genetic elements, prophages, plasmids, and virulence factors that confer adaptive advantages that allow them to cause infections and spread successfully. For example, the clonal expansion of the virulent community-associated MRSA (CA-MRSA) clone USA300 occurred as a result of the acquisition of fluoroquinolone resistance (FQR) and the arginine catabolic mobile element (ACME) (12). The ACME carries genes for arginine catabolism, allowing the bacteria to survive at lower pHs on human skin. In a recent study, we identified four lineage-specific prophages in *S. pseudintermedius* (17). One of these prophages, SpST71A, carries putative antibiotic resistance and virulence genes and is inserted within the *comGA* gene of the genetic competence operon *comG* in the ST71 clone. The *comG* operon in all ST71 isolates is therefore disrupted and likely nonfunctional. The acquisition of SpST71A, together with *comG* disruption, may be contributing to the increased virulence and clonal success of the ST71 clone (17).

Similar to *S. aureus*, a large number of potential virulence factors have been identified in *S. pseudintermedius*, including surface proteins, exoenzymes, and exotoxins (18–21). Many of these virulence factors are encoded by lineage-associated accessory genes likely derived from other bacterial species through HGT. We hypothesize that the successful MDR MRSP clones harbor one or more of these lineage-associated genes that provide them adaptive advantages over other clones. In this study, we have characterized the genome sequences of 622 *S. pseudintermedius* isolates to identify all previously known genes that are likely involved in the virulence and clonal success of the major MDR MRSP clones. We have investigated all putative HGT-derived genes in these strains to gain insights on the lineage-specific acquisition of adaptive functions in the major lineages. A comprehensive analysis of the potential virulence factors and

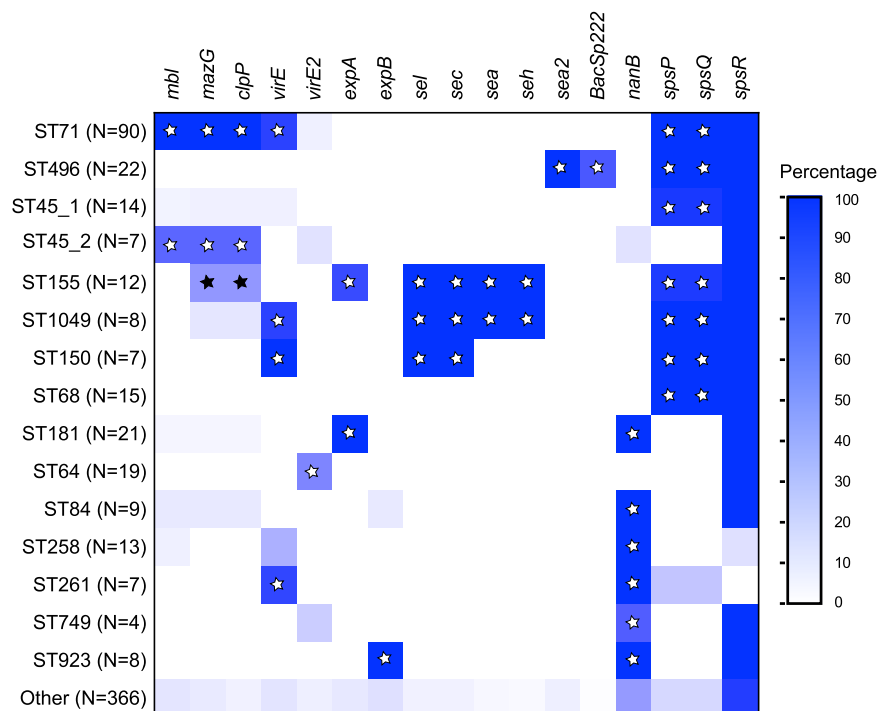


FIG 1 Heatmap showing the distribution of 17 noncore virulence genes. *mbi*, *mazG*, and *clpP* are located on the SpST71A prophage (17). *virE* (locus_tag A9165_10340) and *virE2* (locus_tag EW134_05725) are almost identical in length, with ~76% and ~82% sequence identity at the nucleotide and amino acid levels, respectively. *sea* (locus_tag EW134_02425) and *sea2* (locus_tag EW134_00155) are identical in length, with ~80% and ~72% sequence identity at the nucleotide and amino acid levels, respectively. The *bacSp222* gene, associated with ST496, encodes a 50-amino-acid multifunctional peptide with both bacteriocin and virulence factor characteristics (62). Significant positive associations are indicated with white ($P < 0.005$) and black ($P < 0.05$) stars. N, number of isolates; ST45_1, ST45 sublineage 1; ST45_2, ST45 sublineage 2.

other lineage-associated HGT-derived genes is necessary to better understand the evolution of pathogenicity across different *S. pseudintermedius* lineages. The novel virulence genes identified through comparative genome approaches could serve as targets for therapeutic interventions.

RESULTS AND DISCUSSION

***S. pseudintermedius* harbors lineage-associated virulence genes.** Whole-genome sequencing (WGS) has proven to be a powerful tool in studying the evolution of antibiotic resistance and virulence-associated factors in bacteria. WGS has provided an alternative and complementary high-throughput approach to identifying the underlying genetic factors responsible for clonal expansion of the hypervirulent epidemic clones. Here, we utilized whole-genome sequence data from 622 strains to gain insights on the evolution of pathogenicity and horizontal gene transfer across different *S. pseudintermedius* lineages. Fifty genomes were sequenced as part of this study, while the remaining 572 were retrieved from the NCBI RefSeq or SRA database (Table S1 in the supplemental material). The STs represented by 4 or more isolates were considered predominant and included ST71, ST496, ST45, ST155, ST1049, ST150, ST68, ST181, ST64, ST84, ST258, ST261, ST749, and ST923 ($n = 256$) (Fig. 1). The STs represented by 3 or fewer isolates were considered minor STs ($n = 366$) (Table S1). A total of 69 virulence-associated genes were interrogated in 622 genomes using the large-scale BLAST score ratio (LS-BSR) (Table S2). They belonged to five major functional groups: surface or cell wall-associated proteins, toxins, exoenzymes, proteases, and genetic regulators. Fifty-two of these genes were present in all 622 genomes, irrespective of their sequence types, indicating that they are part of the core genome. The expression levels of the

core virulence genes, however, can vary in different lineages, leading to lineage-wise differences in pathogenicity. The levels of expression of virulence genes are just as important as their lineage-specific presence (22). The frequencies of the remaining 17 virulence-associated genes varied across lineages, and they formed the basis of our virulence gene content analysis (Fig. 1, Table S3).

As shown in Fig. 1, each major lineage is characterized by a unique and specific set of virulence genes. For example, the virus-associated antigen E gene (*virE*), which is carried on a complete prophage, SpST71B, was found in ST71, ST1049, ST150, and ST261 ($P < 0.005$). A novel homologue of *virE*, called *virE2*, was strongly associated with ST64 ($P < 0.005$) (Fig. 1, Fig. S1). Four exfoliative toxin genes, *siet*, *speta*, *expA*, and *expB*, are reported in *S. pseudintermedius* (23–25). ExpA and ExpB toxins have been shown to digest canine desmoglein 1 (Dsg1) and cause intraepidermal splitting in canine skin (24, 25). While *siet* and *speta* genes were found in all 622 strains, *expA* was associated with ST155 and ST181 and *expB* was associated with ST923 ($P < 0.005$) (Fig. 1). Similar distribution patterns of these four toxin-encoding genes have been reported in a previous study (26). The genes encoding staphylococcal enterotoxins (SEs) have been identified frequently in clinical isolates from dogs with cutaneous infections (27). The enterotoxin genes *sel* and *sec*, adjacent to each other in the genome, were associated with ST155, ST1049, and ST150 ($P < 0.005$). The *sea* and *seh* enterotoxin genes, also adjacent to each other, were associated with ST155 and ST1049 ($P < 0.005$). A novel variant of *sea*, called *sea2*, was identified in ST496 (Fig. 1, Fig. S2). The sialidase toxin gene *nanB* was associated with ST181, ST84, ST258, ST261, ST749, and ST923 ($P < 0.005$). Sialidases are important enzymes for sugar acquisition in bacteria, and their role in disease pathogenesis has been shown in many bacterial pathogens (28). In *Streptococcus pneumoniae*, sialidases are involved in colonization and infection of the upper and lower murine respiratory tract (29, 30).

Whole-genome sequencing of *S. pseudintermedius* strain ED99 identified 18 *sps* genes (*spsA* to *spsR*) encoding putative cell wall-associated proteins (CWAPs) (18, 31). The exact functions of only some of these *S. pseudintermedius* surface proteins (Sps) have been deciphered to date. SpsD, SpsL, and SpsO have been shown to participate in bacterial attachment to the host extracellular matrix and, thus, are involved in the pathogenesis of *S. pseudintermedius* (32–34). Similarly, *spsQ*, which is analogous to the *S. aureus* staphylococcal protein A (SpA) gene (*spa*), has been reported to be an important virulence determinant in *S. pseudintermedius* (21, 31). SpsQ (SpA) protein expressed on the cell surface binds to the Fc region of canine IgG, thereby helping the bacteria to escape from opsonization and phagocytosis by the polymorphonuclear cells (21). SpsQ is also a widely used marker for molecular typing (*spa* typing) of *S. pseudintermedius* strains. Unlike in *S. aureus*, there are two orthologues of *spa* in *S. pseudintermedius*, *spsP* (*spa2*) and *spsQ* (*spa1*) (Fig. S3) (31). In ED99 and other *spa*-positive strains, *spsP* and *spsQ* genes are adjacent to each other in the *oriC* environ of the genome, where five other *sps* genes (*spsK*, *spsL*, *spsM*, *spsG*, and *spsJ*) are also located (31). We examined all 18 *sps* genes in our isolates and found that the frequencies of *spsP*, *spsQ* and *spsR* vary among STs (Fig. 1, Table S3). *spsP* and *spsQ* (*spa*) were strongly associated ($P < 0.005$) with ST71, ST496, ST45 sublineage 1, ST155, ST1049, ST150, and ST68 (Fig. 1). Both *spsP* and *spsQ* were absent in ST45 sublineage 2, ST181, ST64, ST84, ST258, ST749, and ST923. The presence or absence of the *spsQ* (*spa*) gene was validated by PCR in 50 *S. pseudintermedius* isolates whose whole genomes were sequenced in this study (Fig. S4). Although previous studies have reported *spa* gene deletion in some *S. pseudintermedius* strains, we have for the first time established a clear association between *spa* gene deletion and genetic lineages (35, 36).

A vaccine solely based on SpA may not be effective against all lineages. Similar to *S. aureus*, the SpA antigen in *S. pseudintermedius* is being considered a potential vaccine candidate antigen (21, 37). However, unlike *S. aureus*, *spa* (*spsQ*) is not a core gene in *S. pseudintermedius* (18, 35). In our globally diverse data set, only 38% of isolates

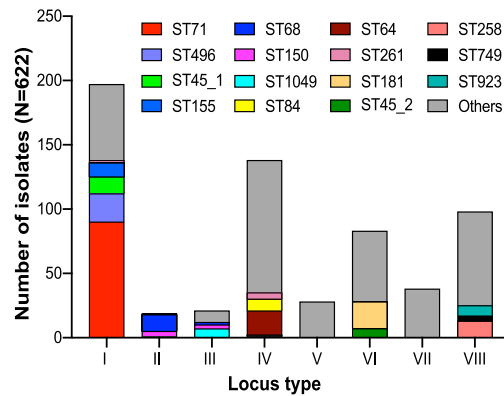
(239 of 622) were *spa* positive, while ~62% of isolates (383 of 622) were *spa* negative, clearly indicating that a large number of strains do not harbor *spa* (Fig. 1, Table S3). It also indicates that *spa* gene function is not essential for bacterial survival, colonization, and establishment of infection. The loss of function resulting from *spa* gene deletion is likely compensated by cell wall-associated proteins encoded by other *sps* genes. It is important to note that the successful MDR MRSP clones ST71, ST68, ST496, and ST45 sublineage 2 have not lost the *spa* gene (Fig. 1). In an important study, Latronico et al. compared the *in vitro* adherence properties of four MRSP (2 ST71:*spa* positive, 1 ST68:*spa* positive, and 1 ST258:*spa* positive) and three MSSP (1 ST25:*spa* positive; 1 ST257:*spa* negative, and 1 ST259:*spa* negative) strains (38). The two *spa*-positive ST71 strains showed greater adherence to canine and human corneocytes than all other *spa*-positive and *spa*-negative non-ST71 MRSP/MSSP strains. The enhanced *in vitro* adherence of ST71 to corneocytes, however, was not linked to *spsQ* (*spa*), *spsP*, or any other CWAP genes examined in the study (38). As mentioned above, the *spa* gene is strongly associated with the genetic background of the strains (Fig. 1). We also examined whether *spa* gene presence is associated with methicillin resistance (*mecA* gene presence). Among 622 isolates analyzed, 323 were MRSP, while 299 were MSSP (Table S3). The *spa* positivity rates in MRSP and MSSP isolates were 56% (181 of 323) and 19% (58 of 299), respectively. Thus, a significant association lies between *spa* and methicillin resistance, where MRSP isolates are more likely to carry *spa* than MSSP isolates ($P < 0.001$). As suggested in previous studies, our results solidify the inefficacy of *spa* typing as a useful method for differentiating all *S. pseudintermedius* strains (15, 16, 35, 39, 40). Instead, *spa* is a useful marker for molecular typing of strains belonging to certain genetic backgrounds. Moreover, a future vaccine consisting solely of SpA antigen may not provide protection against all *S. pseudintermedius* strains, as a large number of them are *spa* negative.

The *spa* region is a hot spot for recombination and horizontal gene transfer.

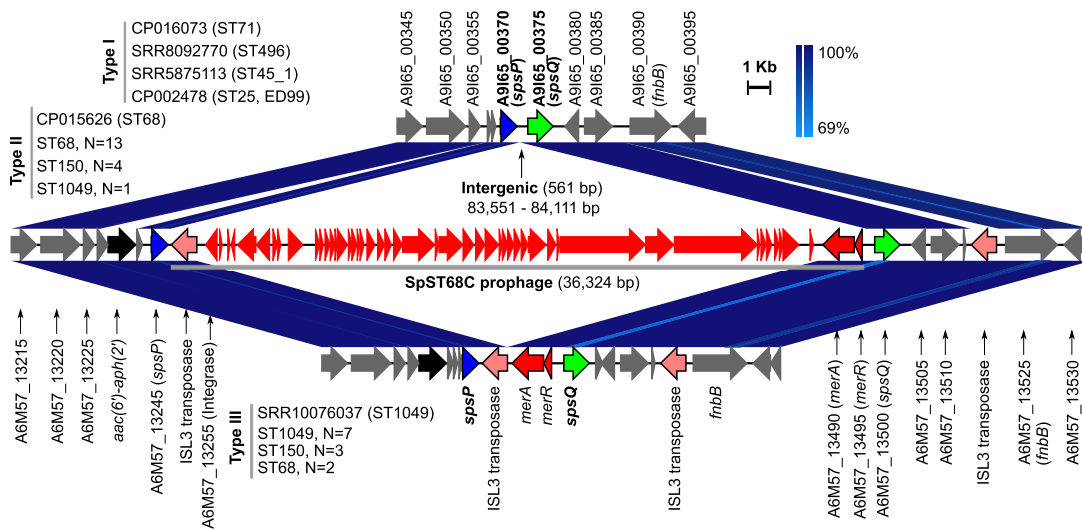
Given that a large number of *S. pseudintermedius* strains do not carry *spa*, it was imperative to investigate the underlying mechanisms of its deletion. Comparative genomic analysis of the *spa* locus (*spsP* and *spsQ* combined) and the flanking sequences revealed that this region is highly variable across strains. The gene contents in this region vary significantly from one lineage to another, suggesting that the *spa* locus is one of the hot spots for recombination and genetic exchange in *S. pseudintermedius*. We classified this region into eight types (locus types I to VIII), based on the genes present in different lineages (Fig. 2A). Locus type I, as found in ST71, ST496, ST155, ST45 sublineage 1, and ST25, contained both *spsP* and *spsQ* genes, adjacent to each other, separated a 561-bp intergenic sequence (Fig. 2B). In ST68, ST150, and ST1049, a mobile genetic element (MGE) carrying putative mercury resistance genes *merA* and *merR* is integrated in the intergenic region between *spsP* and *spsQ* (type II and type III loci). The MGE in locus type II is a complete prophage that we have named SpST68C, as described previously (17). The MGE in locus type III consisted of only *merA* and *merR* genes flanked by ISL3 transposases (Fig. 2B). All *spa*-positive strains ($n = 239$), including minor STs, harbored one of these three locus types (Fig. 2A and B).

Analysis of the corresponding region in *spa*-negative strains ($n = 383$), where both *spsP* and *spsQ* were deleted, identified five different types of loci, types IV, V, VI, VII, and VIII (Fig. 2A and C). In locus type IV, both *spsP* and *spsQ* were deleted, but no new genes were found at the locus where *spa* was deleted. Locus type IV was also the most commonly found locus in minor STs (Fig. 2A). In locus types V and VI, restriction modification (RM) system genes *mbolR*, *dpmM*, and *dpmA* were acquired at the locus where *spa* was deleted (Fig. 2C). The RM system in locus type V had an additional gene that encodes a putative abortive infection phage resistance (AIPR) protein. The AIPR gene is most often found in RM system operons (41). In locus type VII, a set of genes with unknown functions were integrated at the locus where *spa* was deleted. One of these genes was annotated as encoding a CPBP family intramembrane metalloprotease with YdiL (CAAX protease, COG1266, functional category O) and Abi (CAAX protease self-immunity) domains. Abi (abortive infection) proteins are involved in self-immunity

A



B



C

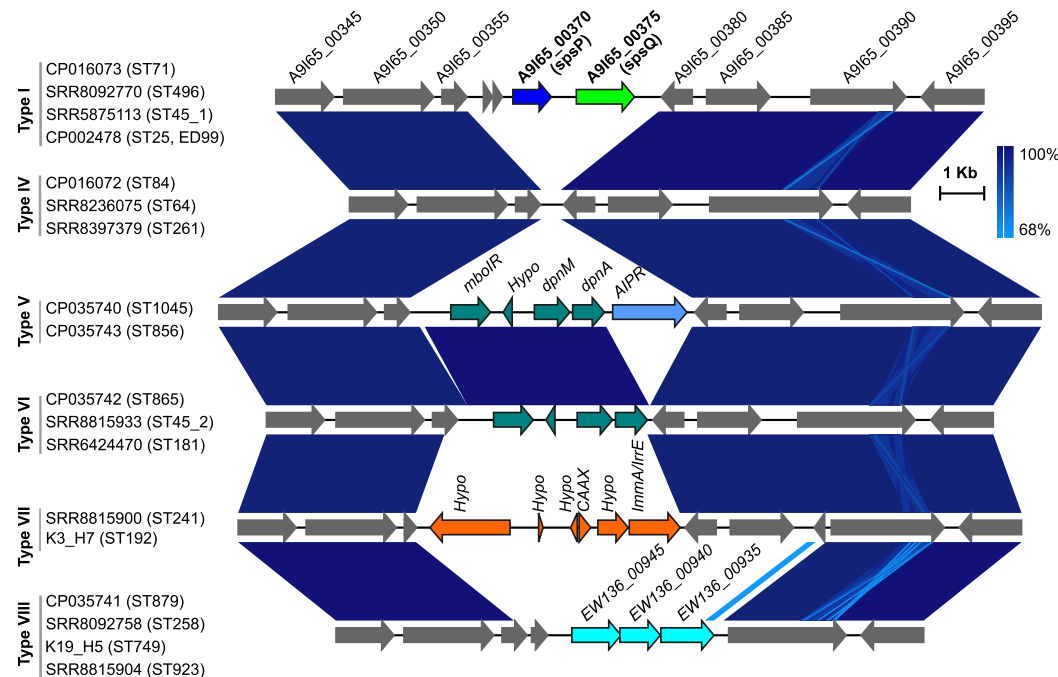


FIG 2 Acquisition of lineage-specific MGEs and accessory genes at the *spa* locus. (A) Eight types of *spa* loci (type I to type VIII) were identified in 622 *S. pseudintermedius* isolates. As shown, the *spa* region is conserved within a lineage (ST) but highly variable across (Continued on next page)

against cognate bacteriocin in some Gram-positive bacteria (42). They have also been shown to provide immunity against bacteriophages by interrupting their DNA replication, transcription, translation, maturation, and lysis steps (43, 44). In locus type VIII, a set of three genes, *EW136_00935* (nucleoside transporter protein, COG1972, functional category F), *EW136_00940* (pseudouridine-5'-phosphate glycosidase, COG2313, functional category F), and *EW136_00945* (winged helix-turn-helix transcriptional regulator), were integrated. Overall, the *spa*-negative strains have lineage-specific genes associated with prophage immunity, barriers to HGT, nucleotide transport, and metabolism at the locus where *spa* was deleted (Fig. 2C). The *spa* locus, therefore, acts as a reservoir where clones have acquired different sets of genes that may be associated with lineage-specific adaptive functions. Further characterization of these genes would be needed to determine their precise roles in *S. pseudintermedius* pathogenesis. These findings also suggest that the deletion of *spa* (*spsP* and *spsQ*) and the integration of new genes at that locus are two independent genetic events. As can be seen in locus type IV, *spa* is deleted but no new genes were acquired at that locus, indicating that *spa* deletion is an independent event. However, it is also possible that the genes were acquired at this locus but may have been lost by genetic drift or purifying selection (45, 46). In locus types II and III, MGEs are integrated without *spsP* and *spsQ* genes being deleted (Fig. 2B).

The horizontal acquisition of genes with adaptive functions has played an important role in the evolution of bacterial pathogens (17, 47). Studies in other bacteria have shown that HGT usually occurs in specific locations, called "hot spots," in the genome (45, 48, 49). Most MGEs and HGT-derived genes linked to antibiotic resistance, metal resistance, and virulence are integrated in the hot spot regions. Furthermore, the virulence genes involved in host cell attachment, invasion, and colonization show higher intensities of recombination and are most often located in the hot spot regions (49). For example, the *tbpA* and *tbpB* genes in *Neisseria gonorrhoeae*, *Neisseria meningitidis*, and *Haemophilus influenzae*, the *slpA* gene in *Clostridium difficile*, and the *clfA*, *sdrC*, *sdrD*, and *sdrE* genes in *S. aureus* are located in the recombination hot spot regions (49). Similar to the *S. pseudintermedius* *spsP* and *spsQ* genes, *tbpA* and *tbpB* in both *Neisseria* and *Haemophilus* are located adjacent to each other and encode surface proteins (49, 50). Collectively, the presence of *spsP* and *spsQ* orthologous genes adjacent to each other in the *oriC* environ and encoding surface proteins that are the target of natural selection makes them potential hot spots for recombination. The ClonalFrameML analysis conducted on a subset of 222 genomes verified the *spa* region as one of the hot spots of recombination in *S. pseudintermedius* (Fig. S5). Apart from ST45, almost all strains within a specific lineage had identical gene contents at the *spa* locus (Fig. 3). Among ST45 isolates, 13 isolates were *spa* positive with the type I locus, while 8 isolates were *spa* negative with the type VI locus. We refer to them as ST45 sublineage 1 (ST45_1) and ST45 sublineage 2 (ST45_2), respectively (Fig. 3).

ST45 is composed of two distinct sublineages with different virulence potential. Twenty-one isolates in our data set were ST45, an MDR MRSP clone widely

FIG 2 Legend (Continued)

lineages (STs). Major STs are represented by different colors, and all minor STs are shown grouped together in gray. (B) Types I, II, and III are *spa*-positive loci. In the type I locus (typified by the sequence with accession number CP016073 [CP016073 type]), *spsP* and *spsQ* genes are adjacent to each other, separated by a 561-bp intergenic region, whereas in the type II locus (CP015626 type), an intact prophage SpST68C (red ORFs), carrying mercury resistance genes *merA* and *merR*, is integrated between *spsP* and *spsQ*. In the type III locus (SRR10076037 type), the SpST68C prophage is deleted, while *merA* and *merR* genes are retained. The representative sequences of ST71 (CP016073), ST68 (CP015626) and ST1049 (SRR10076037) were used to create this linear comparison figure. (C) Type IV, V, VI, VII and VIII loci are *spa* negative. In the type IV locus (CP016072 type), both *spsP* and *spsQ* genes are deleted, but no new genes are integrated at that locus. In type V (CP035740 type) and type VI (CP035742 type) loci, a restriction modification (RM) system operon is integrated at the locus where *spa* was deleted. The RM operon in the type V locus also contains a putative abortive infection phage resistance (AIPR) protein. In the type VII locus (SRR8815900 type), a set of six genes (four hypothetical, one annotated as encoding CAAX protease self-immunity, and one annotated as encoding ImmA/IrrE family proteins) are integrated at the locus where *spa* was deleted. The type VIII locus (CP035741 type) contains three putative genes annotated as encoding NupC/NupG family nucleoside CNT transporter (locus_tag *EW136_00935*), a pseudouridine-5'-phosphate glycosidase (locus_tag *EW136_00940*), and a winged helix-turn-helix transcriptional regulator (locus_tag *EW136_00945*). The *spa*-positive (ST45_1) and *spa*-negative (ST45_2) sublineages of ST45 are also indicated.

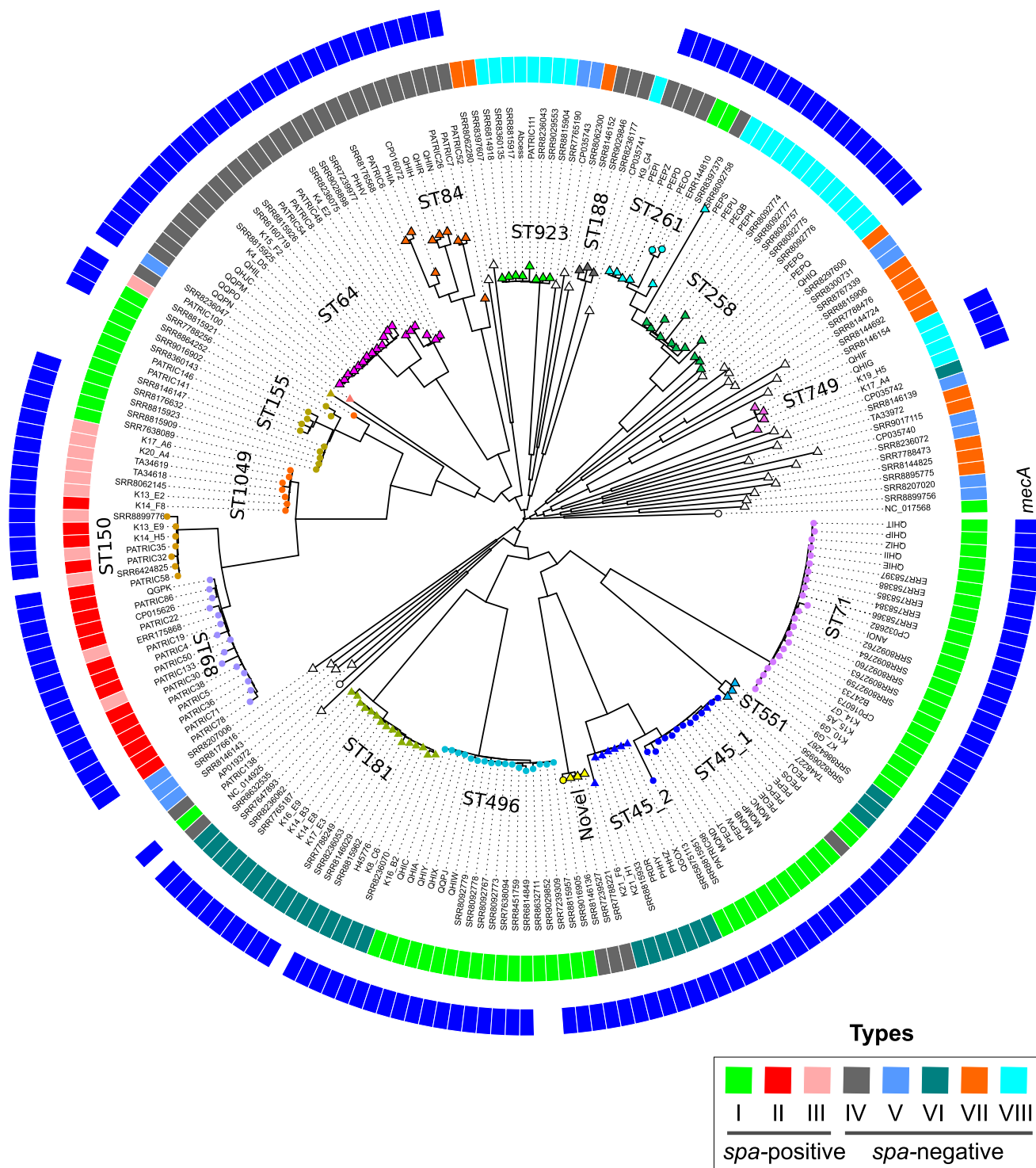


FIG 3 ML phylogeny showing *spa* locus types I to VIII in *S. pseudintermedius* lineages. Only 222 representative strains, including all major lineages and removing redundant STs, were used in this analysis. The genomes were aligned against the reference genome of *S. pseudintermedius* strain 081661 (CP016073), followed by recombination analysis by ClonalFrameML. The recombination-free alignment was used to infer an ML phylogeny using FastTree. Type I to VIII *spa* loci are shown in different colors as indicated. The branch tips labeled with circles are *spa*-positive isolates, while those marked with triangles are *spa*-negative isolates. The outer blue ring indicates *mecA* gene presence (MRSP) in the isolates.

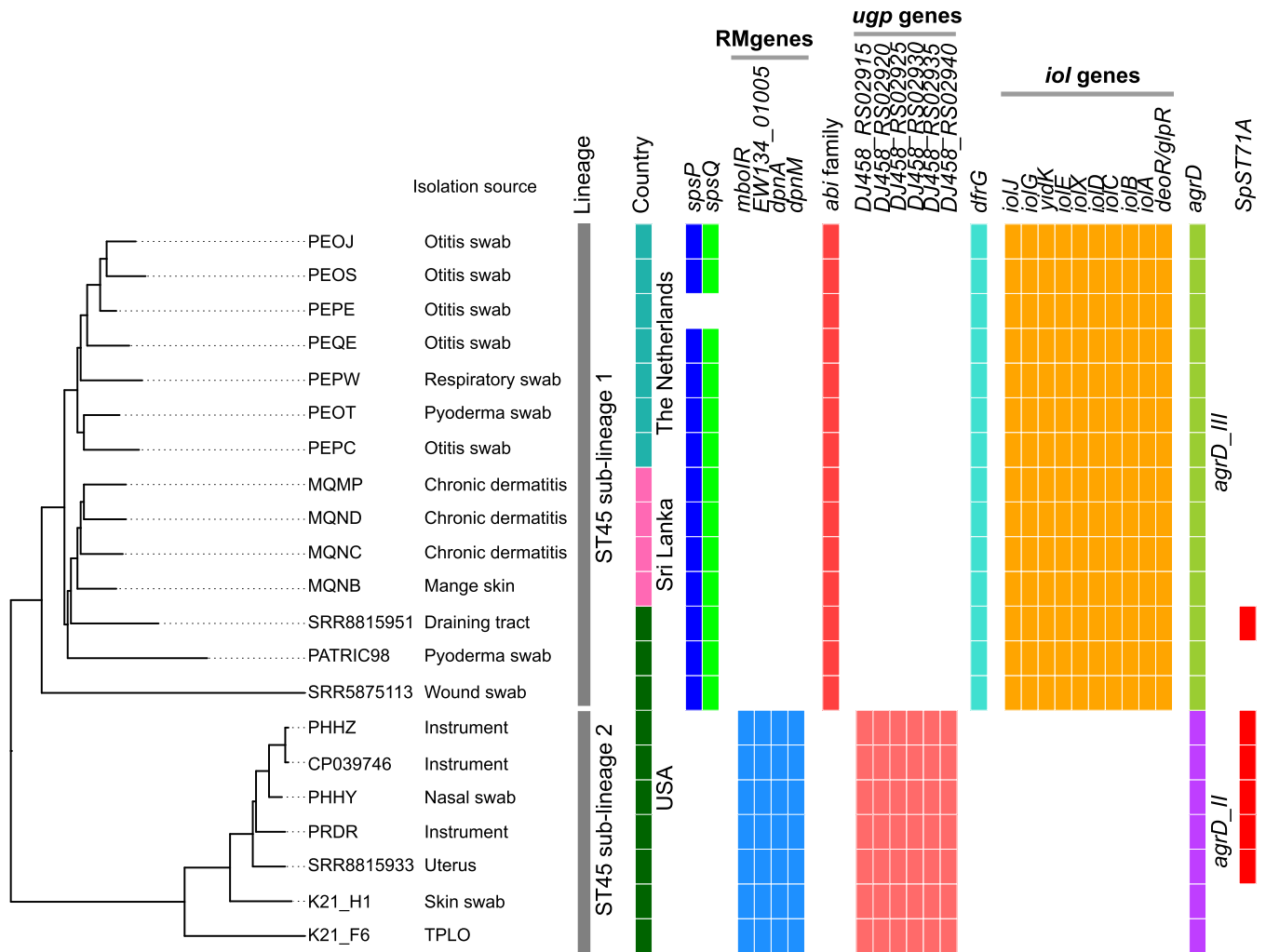


FIG 4 Whole-genome ML phylogeny and pan-GWAS analyses of the ST45 clone. The whole-genome ML phylogeny of the ST45 isolates was created using FastTree. The genes found to be enriched in *spa*-positive sublineage 1 and *spa*-negative sublineage 2 are shown along with the tree. As indicated, ST45 sublineage 1 carries *spsP*, *spsQ*, *dfgG*, *abi*, *iol*, and *agrD* type III genes. ST45 sublineage 2 has lost both *spsP* and *spsQ* genes and has acquired RM system genes *mbolR*, *EW134_01005*, *dpnA*, and *dpnM* at that locus. Similarly, they have lost the *abi* gene and have acquired a putative *ugp* operon (locus_tag *DJ458_RS02915-DJ458_RS02940*) at that locus. All isolates in sublineage 2 carry *agrD* type II, and five of them also harbor the SpST71A prophage (17). Similar to ST71, the SpST71A prophage is inserted within the *comG* operon of these five ST45 isolates, indicating that the prophage might have spread to ST45 from ST71 through HGT. The countries of origin and isolation sources of the isolates are also shown. *DJ458_RS02915* encodes a *sn*-glycerol-3-phosphate ABC transporter; *DJ458_RS02920* encodes a Gfo/Idh/MocA family oxidoreductase; *DJ458_RS02925* encodes a carbohydrate ABC transporter permease; *DJ458_RS02930* encodes a sugar ABC transporter permease; *DJ458_RS02935* encodes a sugar ABC transporter substrate-binding protein; *DJ458_RS02940* encodes a helix-turn-helix domain-containing protein.

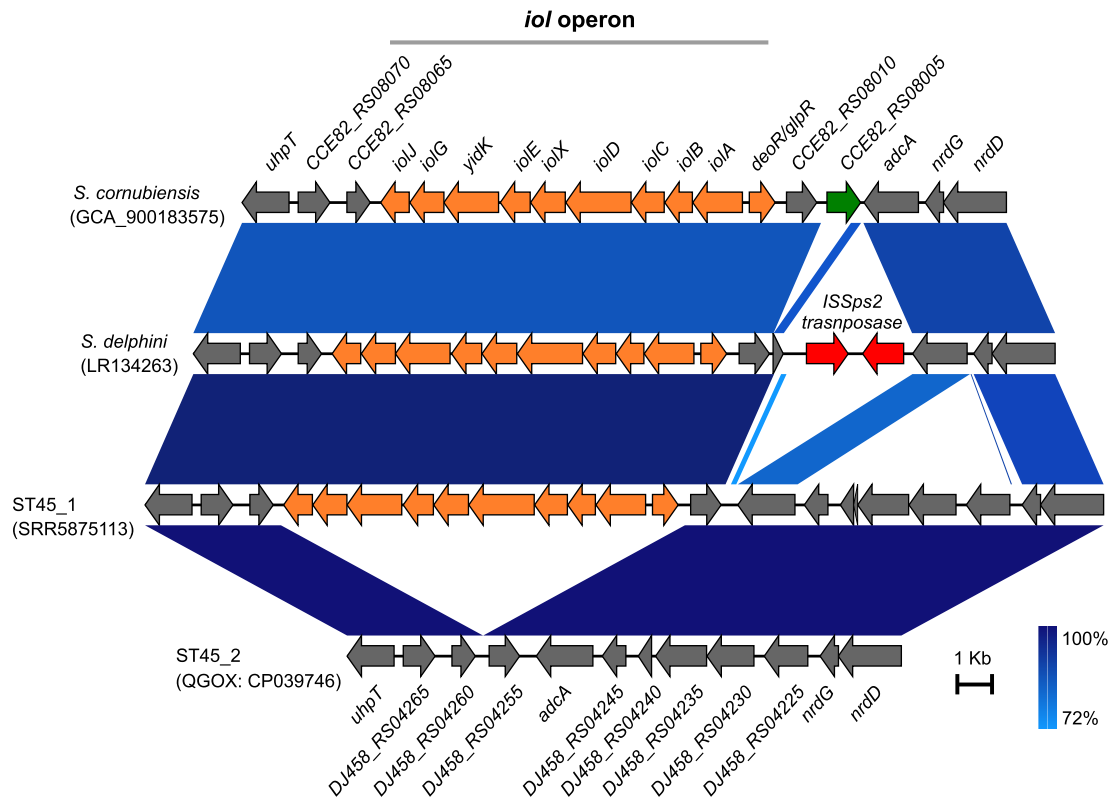
reported from Asia but also reported in the United States and Europe (36). The whole-genome phylogenetic analysis confirmed that ST45 is composed of two distinct sublineages (Fig. 4). ST45 sublineage 1 (ST45_1) contained isolates from Sri Lanka, The Netherlands, and the United States, whereas the sublineage 2 (ST45_2) isolates were all from the United States. To identify other genes that are significantly associated with these two sublineages, we performed pan-GWAS analysis of ST45 isolates using Scoary. The pangenome matrix of ST45 (3,127 total genes, 2,278 core genes, and 849 accessory genes) used for pan-GWAS analysis was constructed using Roary. Besides *spsP* and *spsQ*, several other genes were found to have a strong association (Benjamini-Hochberg-adjusted *P* value of <0.05) with sublineages 1 and 2 (Fig. 4). The RM genes *mbolR*, *dpnM*, and *dpnA* were only associated with sublineage 2, confirming the *spa* results described previously. Since these genes were always acquired at the locus where *spa* was deleted (type VI), no strain simultaneously carried both *spsP* and -Q and RM genes. Additionally, sublineage 2 harbored a putative *ugp* operon consisting of six genes

(locus_tags *DJ458_RS02915-DJ458_RS02940*) associated with *sn*-glycerol-3-phosphate and sugar transport (Fig. 4). The operon was integrated at a locus encoding a putative *abi* family protein in other *S. pseudintermedius* lineages. The *abi* gene and the *ugp* operon were never found simultaneously in any strain, suggesting that the acquisition of the *ugp* operon caused *abi* gene deletion. Pan-GWAS analysis also identified a putative *iol* operon, consisting of *depR-glpR, iolA, iolB, iolC, iolD, iolX, iolE, yidK, iolG, and iolJ* genes, associated with ST45 sublineage 1 (Fig. 4). The *iol* operon has been described in a number of bacterial species, including *Bacillus subtilis*, *Salmonella enterica*, *Clostridium perfringens*, *Staphylococcus xylosus*, and *Lactobacillus casei* (51–54). The bacteria harboring the *iol* operon are capable of utilizing *myo*-inositol (MI) as a carbohydrate source. In *B. subtilis*, the *iol* operon is comprised of *iolABCDEFGHIJ* genes, where *iolG* and *iolE* encode the enzymes required at the first and second steps of MI catabolism, respectively (51). It is likely that ST45 sublineage 1 utilizes MI as an alternative carbon and energy source to colonize and inhabit a particular ecological niche in the animal host or environment. It is important to note that the isolates containing the *iol* operon (ST45_1) were largely from dogs with clinical skin infections, while those with the *ugp* operon (ST45_2) were from medical instruments and opportunistic sites, such as skin and nasal passages, the uterus, and tibial-plateau-leveling osteotomy (TPLO) postoperative infection (Fig. 4). Future experimental studies would be required to establish the lineage-specific roles of *iol* and *ugp* genes in *S. pseudintermedius* pathogenesis. Although both sublineages were multidrug resistant, the *dfrG* gene responsible for trimethoprim resistance was only associated with sublineage 1 (Fig. 4, Table S3). Sublineages 1 and 2 also harbored different alleles of *agrD*, a well-characterized virulence-associated gene in *Staphylococcus*. In a previous study utilizing pulsed-field gel electrophoresis (PFGE), four different clusters were reported within ST45 (36). The cluster B isolates ($n = 2$ Thailand) were *spa* positive, while cluster C ($n = 11$ Thailand), D ($n = 2$ Thailand, $n = 3$ Israel), and E ($n = 2$ Thailand, $n = 10$ Israel) isolates were *spa* negative (36). Furthermore, in cluster B isolates, *spsP* and *spsQ* genes were adjacent to each other, similar to *S. pseudintermedius* strain ED99 (type I in this study), and in cluster C, D, and E isolates, RM genes were integrated at the locus where *spa* was deleted (type VI in this study). While both sublineages have been reported in the United States, only sublineage 1 has been reported in Europe (Fig. 4). Taken together, our results establish that ST45 is composed of two distinct sublineages, both of them having their origin in Asia. They carry different sets of genes related to virulence, prophage immunity, and metabolic pathways and, thus, may have differences in their virulence potentials. A large-scale phylogenetic and GWAS analysis of ST45 isolates from clinical infections and healthy dogs would help to better elucidate the evolution of virulence in these two sublineages.

The lineage-specific genes integrated at the *spa* locus and those identified from the pan-GWAS analysis of ST45 were further examined in all *S. pseudintermedius* genomes and other publicly available genomes of the *Staphylococcus intermedius* group (SIG) (Table S4). Members of the SIG include *S. delphini*, *S. intermedius*, *S. pseudintermedius*, and the newly identified novel species *S. cornubiensis* (20, 55, 56). We discovered that the *iol* and *ugp* operons were also present in all isolates belonging to ST181 ($n = 21$) and some other minor STs of *S. pseudintermedius* (Fig. S6). However, they were not identified in any isolates belonging to other major STs. Most significantly, 13 of the 22 *S. delphini* strains carried the *iol* operon, as does the only reported genome of *S. cornubiensis* (Fig. 5A, Fig. S6). One of the *S. delphini* genomes also carried the *ugp* operon (Fig. 5B, Fig. S6). The *iol* operon in *S. delphini* is located adjacent to a transposable element, suggesting that it may have been acquired from other bacteria through an HGT mechanism.

The *iol* operon and genes acquired at *spa* and *abi* loci are putatively HGT derived. To understand the role of horizontal gene transfer in *S. pseudintermedius* evolution, the putative HGT-derived genes were predicted in each strain (Fig. 6A). Of 10,370 genes estimated in the pangenome of 622 isolates, 771 (~8%) were predicted to be putative HGT-derived genes by the HGTector. This included genes associated with

A



B

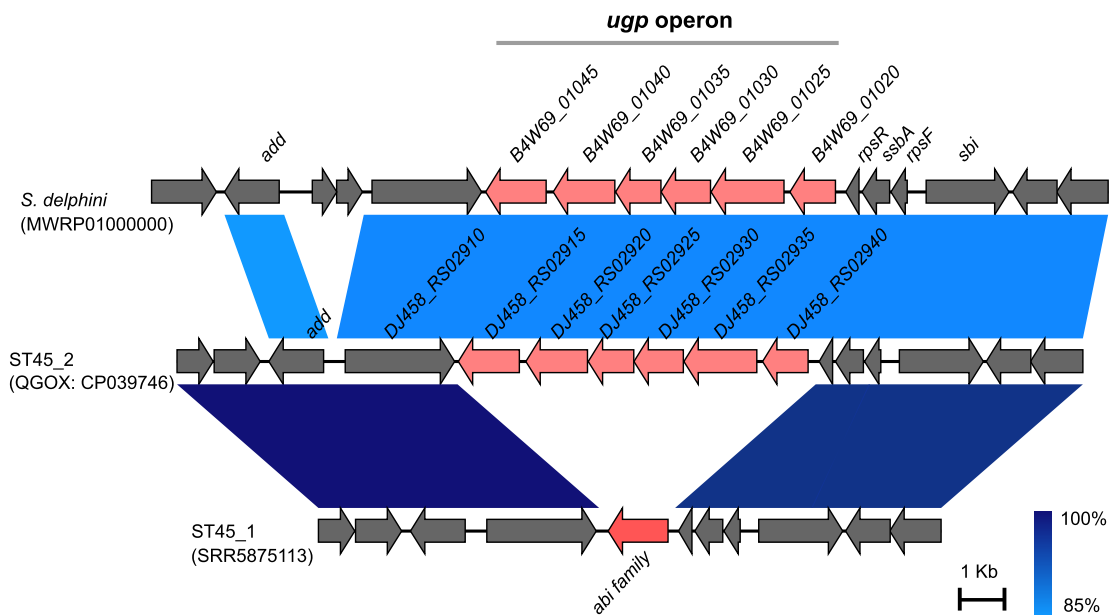


FIG 5 Genomic regions showing the putative *iol* and *ugp* operons in SIG. (A) Alignment showing the *iol* operon (orange ORFs) in *S. cornubiensis*, *S. delphini*, and *S. pseudintermedius* ST45_1. As indicated, the *iol* operon in *S. delphini* is in the vicinity of an insertion sequence (IS) element (ORFs in red). *S. pseudintermedius* ST45_2 does not have an *iol* operon between DJ458_RS04255 and DJ458_RS04260. (B) Alignment showing the putative *ugp* operon (pink ORFs) in *S. delphini* and *S. pseudintermedius* ST45_2. *S. pseudintermedius* ST45_1, on the other hand, has the intact *abi* gene. The accession numbers of the genome sequences used to create these figures are shown in parentheses.

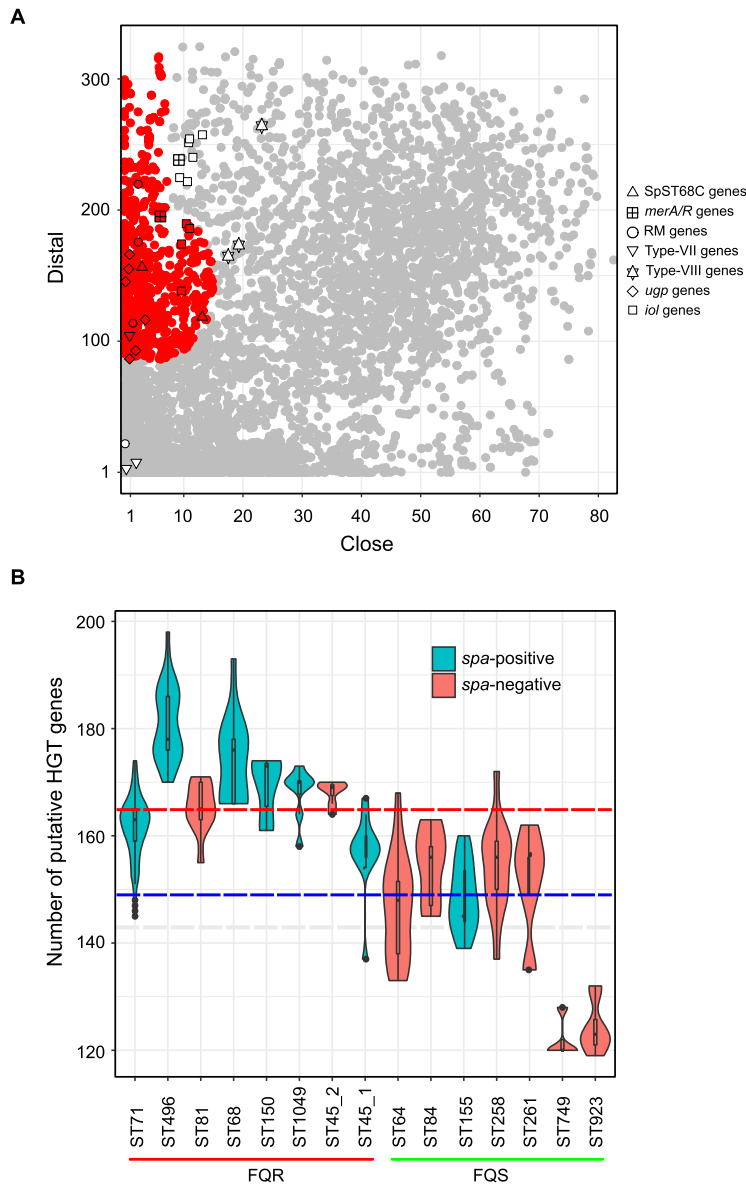


FIG 6 HGT-derived genes in *S. pseudintermedius* pangenome. (A) Scatterplot showing distal and close hits of all 10,370 unique genes (pangenome) from 622 *S. pseudintermedius* strains. The 771 predicted HGT-derived genes are shown in red, while vertically inherited genes (self-group) are shown in gray. As expected, genes associated with antibiotic resistance, virulence, restriction modification, and CRISPR-Cas and genes located on mobile genetic elements, prophages, and plasmids were predicted to be HGT-derived genes. The genes acquired at the *spa* locus and the *abi* locus and *iol* genes are indicated with different shapes. As shown, most of these genes were accurately predicted as HGT derived (within the red color zone), while some of them were close to the HGT group with high distal and low close scores. (B) Violin plot showing numbers of putative HGT-derived genes across *S. pseudintermedius* lineages. For this analysis, the HGT-derived genes were identified individually in each genome. FQR, fluoroquinolone-resistant clones with GyrA Ser84Leu and GrIA Ser80Ile mutations; FQS, fluoroquinolone-sensitive clones with wild-type GyrA and GrIA alleles. The dashed red and blue horizontal lines indicate the median numbers of HGT-derived genes in FQR (median = 165) and FQS (median = 148) clones, respectively.

antibiotic resistance [e.g., *blaZ*, *mecA*, *aac(6')-aph(2'')*, *cat-(pC221)*, *dfrG*, *tet(M)*, and *tet(K)*], plasmids, prophages, transposons, CRISPR-Cas, and RM systems. Most of the genes acquired at the *spa* locus (in types II, III, V, VI, VII, and VIII) were predicted to be HGT-derived genes. Similarly, the genes associated with the ST45_1 and ST45_2 sublineages were also classified as HGT-derived genes. The 771 HGT-derived genes belonged to a wide range of functional categories. A large number of them were linked to DNA replication, recombination, and repair, transcription, defense mechanisms, and

basic metabolic functions. Approximately 20% of them were hypothetical and could not be assigned to any functional category by the eggNOG. We also examined the frequencies of putative HGT-derived genes among the major lineages. As shown in Fig. 6B, ST71, ST496, ST181, ST68, ST150, ST1049, and ST45 carry significantly higher numbers of HGT-derived genes than other clones in our data set. They also carry significantly higher numbers of antibiotic resistance determinants, including the GyrA and GrlA mutations conferring fluoroquinolone resistance (FQR) (17).

In conclusion, (i) the major *S. pseudintermedius* clones carry one or more lineage-associated virulence factors that are likely responsible for their pathogenicity and epidemiological success, (ii) *spa* is not a core gene in *S. pseudintermedius*, since it is deleted in ~62% of strains, (iii) a vaccine based solely on the SpA antigen may not be effective against all strains, (iv) *spa* is a useful marker for molecular typing of only certain *S. pseudintermedius* lineages, (v) the *spa* locus is a hot spot for recombination and horizontal gene transfer, where genes associated with a wide range of functions have been acquired, (vi) ST45 consists of two distinct sublineages with different accessory gene repertoires and, possibly, virulence potential, and (vii) HGT has played an important role in the evolution of *S. pseudintermedius* clones. The lineage-associated genes identified in this study could serve as the basis for future studies on *S. pseudintermedius* pathogenesis and for developing novel therapeutics against this important pathogen.

MATERIALS AND METHODS

Whole-genome sequencing, assembly, and annotation. Whole-genome sequence data of 622 *S. pseudintermedius* isolates representing all predominant STs were analyzed (Table S1 in the supplemental material). Fifty isolates were sequenced as part of this study as previously described (17). Briefly, genomic DNA was isolated from 1 ml of overnight culture grown in tryptic soy broth at 37°C, using the MasterPure Gram-positive DNA purification kit (Lucigen Corp., Middleton, WI). Paired-end library preparation and sequencing (2 × 150-bp paired-end reads, on the Illumina HiSeq 4000) was performed following the standard Illumina protocols (Illumina, Inc., San Diego, CA). Genome assembly was performed using SHOVILL, which utilizes the SPAdes version 2.5.0 assembler at its core (57, 58). The raw sequence reads have been submitted to the National Center for Biotechnology Information (NCBI)'s Sequence Read Archive (SRA) under BioProject accession number [PRJNA564152](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA564152). The remaining 572 were publicly available previously published genomes, retrieved as assemblies from NCBI's RefSeq or as raw sequence reads from the SRA database. These raw reads were also assembled using SHOVILL. All genome assemblies were annotated using Prokka version 1.5.2 (59), and multilocus sequence types (MLSTs) were determined using mlst version 2.18.0 (<https://github.com/tseemann/mlst>).

Screening of virulence genes. We compiled a list of 69 virulence-associated genes from *S. pseudintermedius* and other *Staphylococcus* species (Table S2). This included genes involved in host adaptation, attachment and invasion, host cell lysis, immune evasion, biofilm formation, and regulation of the expression of genes involved in virulence (17, 21, 22, 26, 31, 32, 34, 55, 60–62). The presence of these genes in the genome assemblies of 622 strains was determined using large-scale BLAST score ratio (LS-BSR) with the *blastn* option (63). A gene was considered present with significant similarity if the BSR value was ≥ 0.80 and absent if the BSR was < 0.40 . A BSR value of 0.80 corresponds to ~80% nucleotide identity over 80% of the sequence length (63). The assemblies were also scanned against the virulence factor database (VFDB) to look for additional virulence genes not included in our list (64). Fisher's exact test was used to compare the prevalence of genes among different STs, and a two-sided *P* value of < 0.05 was considered statistically significant.

Analysis of the *spa* locus. As described above, SpsQ (staphylococcal protein A [SpA]) constitutes an important virulence factor and a potential vaccine candidate antigen in *S. pseudintermedius* (21, 65, 66). Additionally, it is a widely used epidemiological marker for molecular typing (*spa* typing) of *S. pseudintermedius* strains. The LS-BSR analysis of the virulence genes described above revealed that the *spa* (*spsQ*) gene is absent in a large number of isolates. We validated the LS-BSR results by alternate methods. We performed *in silico* PCR on all genome assemblies with the widely used *spa*-typing primers S1spaF and S1spaR (14, 35, 36, 67). However, our analysis of the complete *spa* gene sequence alignment from over 200 isolates indicated that the gene in some isolates has polymorphisms in the primer-binding sites and, thus, may not have amplified using the above-named primers. Thus, new primers spaF1 (5'-ACACCAA GTTTCGAGAGAAGGAG-3') and spaR1 (5'-ACTGTTCCACCAGTTGAACGACATG-3') were designed from the highly conserved regions, and another round of *in silico* PCR was performed. The *spa* gene was also verified by *spa* typing of the genomes using *get_spa_type*, a script originally developed for *spa* typing of *S. aureus* (https://github.com/mjsull/spa_typing). The repeat sequences and repeat orders used in this script to call *spa* types were obtained from the Pse-Spa database (<http://www.pse-spa.org>) maintained by Arshnee Moodley.

Analysis of recombination and genetic exchange at the *spa* locus. The *spa*-typeability rate in *S. aureus* is 100%, as the gene is present in all strains, irrespective of their sequence types (68). In contrast, our meta-analysis of the published studies indicated that the *spa*-typeability rate in *S. pseudintermedius*

is ~58%, with studies reporting typeability rates as low as 15 to 20%. Deletion of the *spa* gene in certain *S. pseudintermedius* strains has been reported previously (31, 35). However, no study has investigated the association of *spa* gene deletion with *S. pseudintermedius* genetic lineages. The location of the *spa* gene in the *oriC* region adjacent to an orthologous gene, extensive single nucleotide and repeat polymorphisms, and its role in host-pathogen interaction make it a potential hot spot for recombination. The region around *oriC* in *S. aureus* has been shown to exhibit significantly higher rates of recombination than other regions in the genome (48). We used ClonalFrameML on a select set of 222 genomes, including all major STs, to identify the genome-wide hot spots of recombination in *S. pseudintermedius* (69). The whole-genome alignment used in ClonalFrameML was produced by aligning the genome assemblies against the reference genome of *S. pseudintermedius* strain 081661, using Scapper (<https://github.com/tseemann/scapper>). The maximum-likelihood (ML) phylogeny was inferred from the putative recombination-free alignment using FastTree version 2.1.10 (70). To assess the extent of deletion in the *spa* region, we interrogated the *spa* locus and its flanking regions (6,832 bp upstream and 7,944 bp downstream) in all 622 strains using EasyFig (71). The open reading frames (ORFs) within this region were annotated using eggNOG to get insights on their biological functions (72).

Phylogenetic and pan-GWAS analyses of ST45. We found that only 13 of the 21 ST45 isolates harbored *spa* genes, indicating that there might be two sublineages within ST45 (36). To validate this result, we performed whole-genome-based ML phylogeny and pangenome-wide association (pan-GWAS) analysis of all 21 ST45 isolates (Table S1). Pan-GWAS analysis was performed with Scoary version 1.6.16 (73) on the pangenome matrix generated using Roary version 3.6.8 (17, 74). Scoary was developed to identify the accessory gene(s) associated with a particular phenotypic trait, such as antibiotic resistance, virulence, and group membership. We used the absence (denoted as 0) and presence (denoted as 1) of the *spa* gene as binary traits in Scoary to identify which other genes were enriched in *spa*-positive and *spa*-negative ST45 isolates. Only genes with a Benjamini-Hochberg-adjusted *P* value of <0.05 were considered associated with the presence or absence of *spa*. The genes identified through pan-GWAS were also interrogated in all non-ST45 *S. pseudintermedius* genomes and the publicly available genomes of *S. delphini*, *S. intermedium*, and *S. cornubiensis* (Table S4).

Detection of HGT-derived genes in *S. pseudintermedius*. Genes putatively acquired through horizontal gene transfer (HGT) were identified using HGTector version 2.0b1 (75). The reference pangenome consisting of 10,370 gene clusters predicted from 622 *S. pseudintermedius* strains by Roary, as described above, was used as the input for HGTector. A protein sequence similarity search was performed using Diamond version 0.9.31 (76) against the NCBI nonredundant (nr) database, with minimum search cutoffs set at an E value of $1e-50$, sequence identity of 80%, and percentage query coverage of 80%. The database was created using the default *database* command of HGTector (on 25 March 2020), which downloads, per species, one representative proteome of all microorganisms available in the NCBI RefSeq database. *S. pseudintermedius* (NCBI TaxID 283734) was set as the self group, and genus *Staphylococcus* (NCBI TaxID 1279) as the close group. HGTector analyzes the search results of the proteins and clusters them into the self, close, or distal group according to their NCBI taxonomy. It also calculates a Silhouette score (range of -1 to 1 , with higher values being better) for each gene, a metric which measures the confidence level of a data point assigned to the current cluster in the whole data set. Based on the clustering of “close” and “distal” scores, the program predicts genes that are susceptible to HGT. Genes with low close and high distal scores are considered putative HGT-derived genes.

Data availability. The raw sequence reads of the 50 isolates sequenced as part of this study have been submitted to the NCBI SRA under BioProject accession number [PRJNA564152](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA564152). Accession numbers of all genome sequences used in this study are listed in Table S1.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, EPS file, 0.1 MB.

FIG S2, EPS file, 0.1 MB.

FIG S3, EPS file, 0.1 MB.

FIG S4, EPS file, 0.1 MB.

FIG S5, EPS file, 1.7 MB.

FIG S6, EPS file, 1.3 MB.

TABLE S1, XLSX file, 0.1 MB.

TABLE S2, XLSX file, 0.01 MB.

TABLE S3, XLSX file, 0.1 MB.

TABLE S4, XLSX file, 0.01 MB.

ACKNOWLEDGMENTS

This work was supported by Startup Funds from the College of Veterinary Medicine, UIUC to M.T.A.

We thank Sumiti Vinayak for critical review of this manuscript.

REFERENCES

1. Stegmann R, Burnens A, Maranta CA, Perreten V. 2010. Human infection associated with methicillin-resistant *Staphylococcus pseudintermedius* ST71. *J Antimicrob Chemother* 65:2047–2048. <https://doi.org/10.1093/jac/dkq241>.
2. Guardabassi L, Loeber ME, Jacobson A. 2004. Transmission of multiple antimicrobial-resistant *Staphylococcus intermedius* between dogs affected by deep pyoderma and their owners. *Vet Microbiol* 98:23–27. <https://doi.org/10.1016/j.vetmic.2003.09.021>.
3. Somayaji R, Priyantha MAR, Rubin JE, Church D. 2016. Human infections due to *Staphylococcus pseudintermedius*, an emerging zoonosis of canine origin: report of 24 cases. *Diagn Microbiol Infect Dis* 85:471–476. <https://doi.org/10.1016/j.diagmicrobio.2016.05.008>.
4. Feßler AT, Schuenemann R, Kadlec K, Hensel V, Brombach J, Murugaiyan J, Oechtering G, Burgener IA, Schwarz S. 2018. Methicillin-resistant *Staphylococcus aureus* (MRSA) and methicillin-resistant *Staphylococcus pseudintermedius* (MRSP) among employees and in the environment of a small animal hospital. *Vet Microbiol* 221:153–158. <https://doi.org/10.1016/j.vetmic.2018.06.001>.
5. Van Hoovels L, Vankeerberghen A, Boel A, Van Vaerenbergh K, De Beenhouwer H. 2006. First case of *Staphylococcus pseudintermedius* infection in a human. *J Clin Microbiol* 44:4609–4612. <https://doi.org/10.1128/JCM.01308-06>.
6. Starlander G, Börjesson S, Grönlund Andersson U, Tellgren-Roth C, Melhus A. 2014. Cluster of infections caused by methicillin-resistant *Staphylococcus pseudintermedius* in humans in a tertiary hospital. *J Clin Microbiol* 52:3118–3120. <https://doi.org/10.1128/JCM.00703-14>.
7. Diaz MA, Gardner LB, Libertin CR. 2019. *Staphylococcus pseudintermedius* catheter-related bloodstream infection after exposure to domestic dogs and a cat. *BMJ Case Rep* 12:e231489. <https://doi.org/10.1136/bcr-2019-231489>.
8. Ference EH, Danielian A, Kim HW, Yoo F, Kuan EC, Suh JD. 2019. Zoonotic *Staphylococcus pseudintermedius* sinonasal infections: risk factors and resistance patterns. *Int Forum Allergy Rhinol* 9:724–729. <https://doi.org/10.1002/alar.22329>.
9. Pires Dos Santos T, Damborg P, Moodley A, Guardabassi L. 2016. Systematic review on global epidemiology of methicillin-resistant *Staphylococcus pseudintermedius*: inference of population structure from multilocus sequence typing data. *Front Microbiol* 7:1599. <https://doi.org/10.3389/fmicb.2016.01599>.
10. Moodley A, Damborg P, Nielsen SS. 2014. Antimicrobial resistance in methicillin susceptible and methicillin resistant *Staphylococcus pseudintermedius* of canine origin: literature review from 1980 to 2013. *Vet Microbiol* 171:337–341. <https://doi.org/10.1016/j.vetmic.2014.02.008>.
11. van Duijkeren E, Catry B, Greko C, Moreno MA, Pomba MC, Pyörälä S, Ružauskas M, Sanders P, Threlfall EJ, Torren-Edo J, Törneke K, Scientific Advisory Group on Antimicrobials (SAGAM). 2011. Review on methicillin-resistant *Staphylococcus pseudintermedius*. *J Antimicrob Chemother* 66:2705–2714. <https://doi.org/10.1093/jac/dkr367>.
12. Alam MT, Read TD, Petit RA, Boyle-Vavra S, Miller LG, Eells SJ, Daum RS, David MZ. 2015. Transmission and microevolution of USA300 MRSA in U.S. households: evidence from whole-genome sequencing. *mBio* 6:e00054. <https://doi.org/10.1128/mBio.00054-15>.
13. Tsubakishita S, Kuwahara-Arai K, Sasaki T, Hiramatsu K. 2010. Origin and molecular evolution of the determinant of methicillin resistance in staphylococci. *Antimicrob Agents Chemother* 54:4352–4359. <https://doi.org/10.1128/AAC.00356-10>.
14. Perreten V, Kadlec K, Schwarz S, Grönlund Andersson U, Finn M, Greko C, Moodley A, Kania SA, Frank LA, Bemis DA, Franco A, Iurescia M, Battisti A, Duim B, Wagenaar JA, van Duijkeren E, Weese JS, Fitzgerald JR, Rossano A, Guardabassi L. 2010. Clonal spread of methicillin-resistant *Staphylococcus pseudintermedius* in Europe and North America: an international multicentre study. *J Antimicrob Chemother* 65:1145–1154. <https://doi.org/10.1093/jac/dkq078>.
15. Chanchaithong P, Perreten V, Schwendener S, Tribuddharat C, Chongthaleong A, Niyomtham W, Prapasarakul N. 2014. Strain typing and antimicrobial susceptibility of methicillin-resistant coagulase-positive staphylococcal species in dogs and people associated with dogs in Thailand. *J Appl Microbiol* 117:572–586. <https://doi.org/10.1111/jam.12545>.
16. Kadlec K, Weiß S, Wendlandt S, Schwarz S, Tonpitak W. 2016. Characterization of canine and feline methicillin-resistant *Staphylococcus pseudintermedius* (MRSP) from Thailand. *Vet Microbiol* 194:93–97. <https://doi.org/10.1016/j.vetmic.2016.04.015>.
17. Brooks MR, Padilla-Vélez L, Khan TA, Qureshi AA, Pieper JB, Maddox CW, Alam MT. 2020. Prophage-mediated disruption of genetic competence in *Staphylococcus pseudintermedius*. *mSystems* 5:e00684-19. <https://doi.org/10.1128/mSystems.00684-19>.
18. Bannoehr J, Guardabassi L. 2012. *Staphylococcus pseudintermedius* in the dog: taxonomy, diagnostics, ecology, epidemiology and pathogenicity. *Vet Dermatol* 23:e51–e52. <https://doi.org/10.1111/j.1365-3164.2012.01046.x>.
19. Gonzalez-Martin M, Corbera JA, Suarez-Bonnet A, Tejedor-Junco MT. 2020. Virulence factors in coagulase-positive staphylococci of veterinary interest other than *Staphylococcus aureus*. *Vet Q* 40:118–131. <https://doi.org/10.1080/01652176.2020.1748253>.
20. Fitzgerald JR. 2009. The *Staphylococcus intermedius* group of bacterial pathogens: species re-classification, pathogenesis and the emergence of methicillin resistance. *Vet Dermatol* 20:490–495. <https://doi.org/10.1111/j.1365-3164.2009.00828.x>.
21. Balachandran M, Bemis DA, Kania SA. 2018. Expression and function of protein A in *Staphylococcus pseudintermedius*. *Virulence* 9:390–401. <https://doi.org/10.1080/21505594.2017.1403710>.
22. Couto N, Belas A, Oliveira M, Almeida P, Clemente C, Pomba C. 2016. Comparative RNA-seq-based transcriptome analysis of the virulence characteristics of methicillin-resistant and -susceptible *Staphylococcus pseudintermedius* strains isolated from small animals. *Antimicrob Agents Chemother* 60:e01907-15. <https://doi.org/10.1128/AAC.01907-15>.
23. Futagawa-Saito K, Makino S, Sunaga F, Kato Y, Sakurai-Komada N, Ba-Thein W, Fukuyasu T. 2009. Identification of first exfoliative toxin in *Staphylococcus pseudintermedius*. *FEMS Microbiol Lett* 301:176–180. <https://doi.org/10.1111/j.1574-6968.2009.01823.x>.
24. Iyori K, Hisatsune J, Kawakami T, Shibata S, Murayama N, Ide K, Nagata M, Fukata T, Iwasaki T, Oshima K, Hattori M, Sugai M, Nishifuji K. 2010. Identification of a novel *Staphylococcus pseudintermedius* exfoliative toxin gene and its prevalence in isolates from canines with pyoderma and healthy dogs. *FEMS Microbiol Lett* 312:169–175. <https://doi.org/10.1111/j.1574-6968.2010.02113.x>.
25. Iyori K, Futagawa-Saito K, Hisatsune J, Yamamoto M, Sekiguchi M, Ide K, Son W-G, Olivry T, Sugai M, Fukuyasu T, Iwasaki T, Nishifuji K. 2011. *Staphylococcus pseudintermedius* exfoliative toxin EX1 selectively digests canine desmoglein 1 and causes subcorneal clefts in canine epidermis. *Vet Dermatol* 22:319–326. <https://doi.org/10.1111/j.1365-3164.2011.00952.x>.
26. Little SV, Bryan LK, Hillhouse AE, Cohen ND, Lawhon SD. 2019. Characterization of agr groups of *Staphylococcus pseudintermedius* isolates from dogs in Texas. *mSphere* 4:e00033-19. <https://doi.org/10.1128/mSphere.00033-19>.
27. Yoon JW, Lee G-J, Lee S-Y, Park C, Yoo J-H, Park H-M. 2010. Prevalence of genes for enterotoxins, toxic shock syndrome toxin 1 and exfoliative toxin among clinical isolates of *Staphylococcus pseudintermedius* from canine origin. *Vet Dermatol* 21:484–489. <https://doi.org/10.1111/j.1365-3164.2009.00874.x>.
28. Wang Y-H. 2019. Sialidases from *Clostridium perfringens* and their inhibitors. *Front Cell Infect Microbiol* 9:462. <https://doi.org/10.3389/fcimb.2019.00462>.
29. Manco S, HERNON F, Yesilkaya H, Paton JC, Andrew PW, Kadioglu A. 2006. Pneumococcal neuraminidases A and B both have essential roles during infection of the respiratory tract and sepsis. *Infect Immun* 74:4014–4020. <https://doi.org/10.1128/IAI.01237-05>.
30. Janesch P, Rouha H, Badarau A, Stulik L, Mirkina I, Caccamo M, Havlicek K, Maierhofer B, Weber S, Groß K, Steinhäuser J, Zerbs M, Varga C, Dolezilkoval I, Maier S, Zauner G, Nielsen N, Power CA, Nagy E. 2018. Assessing the function of pneumococcal neuraminidases NanA, NanB and NanC in in vitro and in vivo lung infection models using monoclonal antibodies. *Virulence* 9:1521–1538. <https://doi.org/10.1080/21505594.2018.1520545>.
31. Bannoehr J, Ben Zakour NL, Reglinski M, Inglis NF, Prabhakaran S, Fossum E, Smith DG, Wilson GJ, Cartwright RA, Haas J, Hook M, van den Broek AHM, Thoday KL, Fitzgerald JR. 2011. Genomic and surface proteomic analysis of the canine pathogen *Staphylococcus pseudinterme-*

- dict reveals proteins that mediate adherence to the extracellular matrix. *Infect Immun* 79:3074–3086. <https://doi.org/10.1128/IAI.00137-11>.
32. Bannoehr J, Brown JK, Shaw DJ, Fitzgerald RJ, van den Broek AHM, Thoday KL. 2012. Staphylococcus pseudintermedius surface proteins SpsD and SpsO mediate adherence to ex vivo canine corneocytes. *Vet Dermatol* 23:e26. <https://doi.org/10.1111/j.1365-3164.2011.01021.x>.
 33. Richards AC, O'Shea M, Beard PM, Goncheva MI, Tuffs SW, Fitzgerald JR, Lengeling A. 2018. Staphylococcus pseudintermedius surface protein L (SpsL) is required for abscess formation in a Murine model of cutaneous infection. *Infect Immun* 86:e00631-18. <https://doi.org/10.1128/IAI.00631-18>.
 34. Pietrocchia G, Gianotti V, Richards A, Nobile G, Geoghegan JA, Rindi S, Monk IR, Bordt AS, Foster TJ, Fitzgerald JR, Speziale P. 2015. Fibronectin binding proteins SpsD and SpsL both support invasion of canine epithelial cells by Staphylococcus pseudintermedius. *Infect Immun* 83:4093–4102. <https://doi.org/10.1128/IAI.00542-15>.
 35. Moodley A, Stegger M, Ben Zakour NL, Fitzgerald JR, Guardabassi L. 2009. Tandem repeat sequence analysis of staphylococcal protein A (spa) gene in methicillin-resistant Staphylococcus pseudintermedius. *Vet Microbiol* 135:320–326. <https://doi.org/10.1016/j.vetmic.2008.09.070>.
 36. Perreten V, Chanchaithong P, Prapasarakul N, Rossano A, Blum SE, Elad D, Schwendener S. 2013. Novel pseudo-staphylococcal cassette chromosome mec element (ψ SCCmec57395) in methicillin-resistant Staphylococcus pseudintermedius CC45. *Antimicrob Agents Chemother* 57:5509–5515. <https://doi.org/10.1128/AAC.00738-13>.
 37. Kim HK, Cheng AG, Kim H-Y, Missiakos DM, Schneewind O. 2010. Non-toxicogenic protein A vaccine for methicillin-resistant Staphylococcus aureus infections in mice. *J Exp Med* 207:1863–1870. <https://doi.org/10.1084/jem.20092514>.
 38. Latronico F, Moodley A, Nielsen S, Guardabassi L. 2014. Enhanced adherence of methicillin-resistant Staphylococcus pseudintermedius sequence type 71 to canine and human corneocytes. *Vet Res* 45:70–77. <https://doi.org/10.1186/1297-9716-45-70>.
 39. Kang J-H, Chung T-H, Hwang C-Y. 2017. Clonal distribution of methicillin-resistant Staphylococcus pseudintermedius isolates from skin infection of dogs in Korea. *Vet Microbiol* 210:32–37. <https://doi.org/10.1016/j.vetmic.2017.08.017>.
 40. Ishihara K, Koizumi A, Saito M, Muramatsu Y, Tamura Y. 2016. Detection of methicillin-resistant Staphylococcus pseudintermedius ST169 and novel ST354 SCCmec II-III isolates related to the worldwide ST71 clone. *Epidemiol Infect* 144:434–442. <https://doi.org/10.1017/S0950268815001545>.
 41. Iyer LM, Abhiman S, Aravind L. 2008. MutL homologs in restriction-modification systems and the origin of eukaryotic MORC ATPases. *Biol Direct* 3:8–9. <https://doi.org/10.1186/1745-6150-3-8>.
 42. Kjos M, Snipen L, Salehian Z, Nes IF, Diep DB. 2010. The abi proteins and their involvement in bacteriocin self-immunity. *J Bacteriol* 192:2068–2076. <https://doi.org/10.1128/JB.01553-09>.
 43. Samson JE, Belanger M, Moineau S. 2013. Effect of the abortive infection mechanism and type III toxin/antitoxin system AbiQ on the lytic cycle of Lactococcus lactis phages. *J Bacteriol* 195:3947–3956. <https://doi.org/10.1128/JB.00296-13>.
 44. Chopin M-C, Chopin A, Bidnenko E. 2005. Phage abortive infection in lactococci: variations on a theme. *Curr Opin Microbiol* 8:473–479. <https://doi.org/10.1016/j.mib.2005.06.006>.
 45. Oliveira PH, Touchon M, Cury J, Eduardo PCR. 2020. The chromosomal organization of horizontal gene transfer in bacteria. *Nat Commun* 11:1–11. <https://doi.org/10.1038/s41467-020-15091-5>.
 46. Koskiniemi S, Sun S, Berg OG, Andersson DI. 2012. Selection-driven gene loss in bacteria. *PLoS Genet* 8:e1002787. <https://doi.org/10.1371/journal.pgen.1002787>.
 47. McCarthy AJ, Loeffler A, Witney AA, Gould KA, Lloyd DH, Lindsay JA. 2014. Extensive horizontal gene transfer during Staphylococcus aureus co-colonization in vivo. *Genome Biol Evol* 6:2697–2708. <https://doi.org/10.1093/gbe/evu214>.
 48. Everitt RG, Didelot X, Batty EM, Miller RR, Knox K, Young BC, Bowden R, Auton A, Votintseva A, Larner-Svensson H, Charlesworth J, Golubchik T, Ip CLC, Godwin H, Fung R, Peto TEA, Walker AS, Crook DW, Wilson DJ. 2014. Mobile elements drive recombination hotspots in the core genome of Staphylococcus aureus. *Nat Commun* 5:3956. <https://doi.org/10.1038/ncomms4956>.
 49. Yahara K, Didelot X, Jolley KA, Kobayashi I, Maiden MCJ, Sheppard SK, Falush D. 2016. The landscape of realized homologous recombination in pathogenic bacteria. *Mol Biol Evol* 33:456–471. <https://doi.org/10.1093/molbev/msv237>.
 50. Schryvers AB, Stojiljkovic I. 1999. Iron acquisition systems in the pathogenic Neisseria. *Mol Microbiol* 32:1117–1123. <https://doi.org/10.1046/j.1365-2958.1999.01411.x>.
 51. Yoshida K-I, Yamaguchi M, Morinaga T, Kinehara M, Ikeuchi M, Ashida H, Fujita Y. 2008. myo-Inositol catabolism in Bacillus subtilis. *J Biol Chem* 283:10415–10424. <https://doi.org/10.1074/jbc.M708043200>.
 52. Kröger C, Fuchs TM. 2009. Characterization of the myo-inositol utilization island of Salmonella enterica serovar Typhimurium. *J Bacteriol* 191:545–554. <https://doi.org/10.1128/JB.01253-08>.
 53. Dordet-Frisoni E, Dorchies G, De Araujo C, Talon R, Leroy S. 2007. Genomic diversity in Staphylococcus xylosum. *Appl Environ Microbiol* 73:7199–7209. <https://doi.org/10.1128/AEM.01629-07>.
 54. Yebra MJ, Zúñiga M, Beaufils S, Pérez-Martínez G, Deutscher J, Monedero V. 2007. Identification of a gene cluster enabling Lactobacillus casei BL23 to utilize myo-inositol. *Appl Environ Microbiol* 73:3850–3858. <https://doi.org/10.1128/AEM.00243-07>.
 55. Ben Zakour NL, Beatson SA, van den Broek AHM, Thoday KL, Fitzgerald JR. 2012. Comparative genomics of the Staphylococcus intermedius group of animal pathogens. *Front Cell Infect Microbiol* 2:44. <https://doi.org/10.3389/fcimb.2012.00044>.
 56. Murray AK, Lee J, Bendall R, Zhang L, Sunde M, Schau Slettemeås J, Gaze W, Page AJ, Vos M. 2018. Staphylococcus cornubiensis sp. nov., a member of the Staphylococcus intermedius group (SIG). *Int J Syst Evol Microbiol* 68:3404–3408. <https://doi.org/10.1099/ijsem.0.002992>.
 57. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirtkin 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>.
 58. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
 59. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
 60. O'Gara JP. 2007. ica and beyond: biofilm mechanisms and regulation in Staphylococcus epidermidis and Staphylococcus aureus. *FEMS Microbiol Lett* 270:179–188. <https://doi.org/10.1111/j.1574-6968.2007.00688.x>.
 61. Bannoehr J, Ben Zakour NL, Waller AS, Guardabassi L, Thoday KL, van den Broek AHM, Fitzgerald JR. 2007. Population genetic structure of the Staphylococcus intermedius group: insights into agr diversification and the emergence of methicillin-resistant strains. *J Bacteriol* 189:8685–8692. <https://doi.org/10.1128/JB.01150-07>.
 62. Władysław B, Piejko M, Bzowska M, Pieta P, Krzysik M, Mazurek Ł, Guevara-Lora I, Bukowski M, Sabat AJ, Friedrich AW, Bonar E, Miedzobrodzki J, Dubin A, Mak P. 2015. A peptide factor secreted by Staphylococcus pseudintermedius exhibits properties of both bacteriocins and virulence factors. *Sci Rep* 5:14569. <https://doi.org/10.1038/srep14569>.
 63. Sahl JW, Sistrunk JR, Fraser CM, Hine E, Baby N, Begum Y, Luo Q, Sheikh A, Qadri F, Fleckenstein JM, Rasko DA. 2015. Examination of the enterotoxigenic Escherichia coli population structure during human infection. *mBio* 6:e00501-15. <https://doi.org/10.1128/mBio.00501-15>.
 64. Liu B, Zheng D, Jin Q, Chen L, Yang J. 2019. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res* 47:D687–D692. <https://doi.org/10.1093/nar/gky1080>.
 65. Grandolfo E. 2018. Looking through Staphylococcus pseudintermedius infections: could SpA be considered a possible vaccine target? *Virulence* 9:703–706. <https://doi.org/10.1080/21505594.2018.1426964>.
 66. Abouelkhair MA, Bemis DA, Kania SA. 2018. Characterization of recombinant wild-type and nontoxicogenic protein A from Staphylococcus pseudintermedius. *Virulence* 9:1050–1061. <https://doi.org/10.1080/21505594.2018.1489199>.
 67. Ruscher C, Lübke-Becker A, Semmler T, Wlekinski C-G, Paasch A, Šoba A, Stamm I, Kopp P, Wieler LH, Walther B. 2010. Widespread rapid emergence of a distinct methicillin- and multidrug-resistant Staphylococcus pseudintermedius (MRSP) genetic lineage in Europe. *Vet Microbiol* 144:340–346. <https://doi.org/10.1016/j.vetmic.2010.01.008>.
 68. Hallin M, Friedrich AW, Struelens MJ. 2009. spa typing for epidemiological surveillance of Staphylococcus aureus. *Methods Mol Biol* 551:189–202. https://doi.org/10.1007/978-1-60327-999-4_15.
 69. Didelot X, Wilson DJ. 2015. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Comput Biol* 11:e1004041. <https://doi.org/10.1371/journal.pcbi.1004041>.
 70. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>.
 71. Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome compar-

- ison visualizer. *Bioinformatics* 27:1009–1010. <https://doi.org/10.1093/bioinformatics/btr039>.
72. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, Mering von C, Bork P. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* 44:D286–D293. <https://doi.org/10.1093/nar/gkv1248>.
 73. Brynildsrud O, Bohlin J, Scheffer L, Eldholm V. 2016. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biol* 17:238–239. <https://doi.org/10.1186/s13059-016-1108-8>.
 74. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>.
 75. Zhu Q, Kosoy M, Dittmar K. 2014. HGTector: an automated method facilitating genome-wide discovery of putative horizontal gene transfers. *BMC Genomics* 15:717–718. <https://doi.org/10.1186/1471-2164-15-717>.
 76. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>.