



OPEN

Shared antibiotic resistance and virulence genes in *Staphylococcus aureus* from diverse animal hosts

Spencer A. Bruce^{1✉}, Joshua T. Smith^{2,3}, Jennifer L. Mydosh², John Ball², David B. Needle^{2,4}, Robert Gibson⁴ & Cheryl P. Andam^{1✉}

The emergence of methicillin-resistant *Staphylococcus aureus* (MRSA) poses an important threat in human and animal health. In this study, we ask whether resistance and virulence genes in *S. aureus* are homogeneously distributed or constrained by different animal hosts. We carried out whole genome sequencing of 114 *S. aureus* isolates from ten species of animals sampled from four New England states (USA) in 2017–2019. The majority of the isolates came from cats, cows and dogs. The maximum likelihood phylogenetic tree based on the alignment of 89,143 single nucleotide polymorphisms of 1173 core genes reveal 31 sequence types (STs). The most common STs were ST5, ST8, ST30, ST133 and ST2187. Every genome carried at least eight acquired resistance genes. Genes related to resistance found in all genomes included *norA* (fluoroquinolone), *arlRS* (fluoroquinolone), *lmrS* (multidrug), *tet(38)* (tetracycline) and *mepAR* (multidrug and tigecycline resistance). The most common superantigen genes were *tsst-1*, *sea* and *sec*. Acquired antibiotic resistance ($n=10$) and superantigen ($n=9$) genes of *S. aureus* were widely shared between *S. aureus* lineages and between strains from different animal hosts. These analyses provide insights for considering bacterial gene sharing when developing strategies to combat the emergence of high-risk clones in animals.

Staphylococcus aureus is notable for its ability to colonize a wide range of vertebrate hosts, with each host representing a distinct ecological niche. In humans, it is responsible for severe nosocomial and community-associated infections^{1–3}. *S. aureus* infections in animals are most commonly reported as a cause of mastitis in dairy-producing animals (including cattle, sheep and goats), osteomyelitis in poultry, and skin abscesses, mastitis and septicemia in farmed rabbits⁴. *S. aureus* has been reported in companion animals (e.g., dogs, cats)^{5,6} and wildlife^{7,8}. Animal-associated *S. aureus* is also known to carry diverse antibiotic resistance genes and can therefore act as a reservoir for transmission of resistant strains to humans⁹. The emergence and spread of multidrug-resistant and methicillin-resistant *S. aureus* (MRSA) are therefore not limited to hospital and community settings, but also pose an important threat in veterinary medicine, agricultural systems and food production. Therefore, understanding the population and genome dynamics of MRSA across multiple eukaryotic hosts is key to advancing a One Health approach to attain optimal health for people, domestic animals, wildlife, plants, and the environment¹⁰.

The ability to switch and adapt to new hosts are a dynamic and a constant feature of the evolution of *S. aureus*. Some host transitions are particularly notable. Two lineages in the clonal complex (CC) 97 that cause epidemic community-associated MRSA in humans was inferred to have evolved from independent bovine-to-human host jumps more than 40 years ago¹¹. Methicillin resistance in this clone was acquired after the host-jump, likely as a result of strong selective pressures imposed by the widespread use of antibiotics for treating human infections¹¹. Host-jumping events can also happen from humans to animals, as in the case of sequence type 5 (ST5) in broiler chickens that jumped from humans in the early 1970s¹². Another lineage, CC398, has jumped from humans to livestock and back to humans, with resistance to methicillin and tetracycline acquired after the introduction to livestock from humans¹³. The majority of healthcare-associated infections of livestock-associated MRSA CC398 in Denmark from 2014 to 2016 were shown to be the result of repeated spillover introductions from

¹Department of Biological Sciences, University at Albany, State University of New York, Albany, NY 12222, USA. ²Department of Molecular, Cellular and Biomedical Sciences, University of New Hampshire, Durham, NH 03824, USA. ³Infectious Disease and Microbiome Program, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. ⁴New Hampshire Veterinary Diagnostic Laboratory, Durham, NH 03824, USA. ✉email: sbruce@albany.edu; candam@albany.edu

nearby pig farms into healthcare facilities¹⁴. Human-to-monkey jumps of *S. aureus* ST6 and ST15 that occurred 2700 years ago was facilitated by the loss of phage-carrying genes that played a role in human colonization¹⁵. Over the past 5000–6000 years, numerous instances of cross-species transmission events have occurred in *S. aureus*¹⁶. These host-species transitions have been associated with horizontal gene transfer (HGT) of genetic elements conferring traits required for survival in the new host¹⁶. Some of these cross-species jumps have resulted in the emergence of successful outbreak-causing and widely circulating *S. aureus* lineages in humans and animals¹⁶.

The generalist nature of a pathogen can provide many opportunities for mobilizable DNA to be disseminated more widely among strains that colonize different hosts. However, it remains unclear whether for multi-host bacteria such as *S. aureus*, horizontally transferred genes follow the distribution of the bacterial cell that harbor them or create their own distribution. In this study, we sought to elucidate the population genomic structure of clinical animal-associated *S. aureus*. We ask whether the gene pool of resistance and virulence in *S. aureus* is homogeneously distributed or constrained by different animal hosts. Using whole genome sequencing of 114 *S. aureus* isolates from ten animal species, we demonstrate that a remarkable array of acquired resistance genes and superantigen genes of *S. aureus* are widely shared between multiple *S. aureus* lineages and between strains from different animal hosts. This study provides important insights for considering bacterial gene sharing between different animal hosts in developing strategies to combat the emergence of high-risk clones in animals.

Results

Phylogenetic diversity of animal-associated *S. aureus* in New England. We obtained a total of 114 high quality draft genomes of *S. aureus* isolates obtained through routine diagnostic tests of clinical specimens from diseased animals. These were submitted to the New Hampshire Veterinary Diagnostic Laboratory (NHVDL) from October 2017 to October 2019 (Fig. 1A and Supplementary Table S1). We obtained the isolates from four states in the United States: New Hampshire ($n=74$), Maine ($n=13$), Massachusetts ($n=12$) and Vermont ($n=15$) (Fig. 1B). The majority of isolates came from cows ($n=30$), dogs ($n=28$) and cats ($n=25$) (Fig. 1C). Other domestic animals from which we obtained isolates included horses, goats and rabbits, while wild animals included birds, rodents, deer, rabbits and a tortoise.

De novo assembly of the 114 genomes generated sequences of sizes ranging from 2.68 to 2.90 Mb (mean = 2.77 Mb). The number of predicted genes ranged from 2429 to 2739 per genome (mean = 2570) (Supplementary Table S1). The pan-genome of the New England *S. aureus* population consisted of 7500 orthologous gene families. The genes in the pan-genome were categorized into core genes (present in 99% of genomes), soft-core genes (present in 95% to < 99% of genomes), shell genes (present in 15% to < 95% of genomes), and cloud genes (present in < 15% of genomes). We identified 1773 core genes (present in 113–114 genomes), 116 soft core genes (present in 109–112 genomes), 1158 shell genes (present in 18–108 genomes) and 4453 cloud genes (present in 1–17 genomes) (Supplementary Table S2). The combined core and soft-core genes comprised 25.19% of the pan-genome, while the combined shell and cloud genes (which together make up the accessory genome) comprised 74.81% of the pan-genome. We identified 1716 genes, representing 22.88% of the species pan-genome, that are unique to a single strain.

The maximum likelihood phylogenetic tree based on the alignment of 89,143 SNPs of the core genes revealed many deep branching lineages consisting of 31 previously identified STs (Fig. 1A). The most common STs were ST5 ($n=15$ genomes), ST8 ($n=7$ genomes), ST30 ($n=8$ genomes), ST133 ($n=7$ genomes) and ST2187 ($n=16$ genomes). Despite the unequal number of isolates from the four states (Fig. 1B), the phylogeny showed a lack of structure relative to the state from which the isolate originated.

Some STs were obtained from multiple animal host species (Fig. 1C), indicating a broad host range: ST5 in cat ($n=3$), dog ($n=8$), bird ($n=4$); ST8 in cat ($n=4$), dog ($n=3$); ST30 in cat ($n=3$), dog ($n=3$), horse ($n=1$), reptile ($n=1$); ST133 in horse ($n=3$), dog ($n=1$), goat ($n=1$), rabbit ($n=1$), cow ($n=1$); ST398 in horse ($n=3$), bird ($n=1$), cow ($n=1$), goat ($n=1$). STs of which all isolates were associated with only a single animal species included STs 151, 352 and 2187, which all originated from cows. However, this ST distribution among animal hosts may reflect the disproportionate number of isolates from cattle in our dataset and is not necessarily reflective of host-restricted niches.

Distribution of acquired genes related to antibiotic resistance. We determined the presence of acquired genes associated with resistance to different antibiotics. These genes represent a variety of resistance mechanisms (drug inactivation, target alteration, efflux, target replacement, target protection) based on definitions in the CARD database¹⁷. We identified a total of 30 resistance genes across the entire dataset (Fig. 2A and Supplementary Table S3).

In silico detection of the *mecA* gene from the genome sequences revealed seven MRSA isolates and 107 methicillin-susceptible *S. aureus* (MSSA) isolates. The *mecA* gene encodes an extra penicillin-binding protein (PBP2a) that has low affinity to virtually all beta-lactam antibiotics^{18,19}. The *mecA* gene is carried by a mobile chromosomal cassette *SCCmec* and is classified into types based on the combination of the *ccr* and *mec* complexes they carry¹⁹. To date, there are 14 structurally distinct *SCCmec* types (I–XIV) that have been described^{19–21}. *SCCmec* typing of the New England *S. aureus* population revealed the presence of types II ($n=4$ genomes), III ($n=1$ genomes) and IV ($n=2$ genomes) (Fig. 2A and Supplementary Table S1). *SCCmec* type II was detected in ST 5, type III in ST1 and type IV in STs 8 and 72. We did not detect the gene *mecC*, which is a divergent form of *mecA* and also mediates beta-lactam resistance²². We found a slight discrepancy between the in silico detection of the *mecA* gene and in vitro phenotypic testing for methicillin resistance. There were two isolates (3659B and 4052C) whose genomes did not contain the *mecA* gene but were phenotypically tested as MRSA. These alternative mechanisms warrant further study.

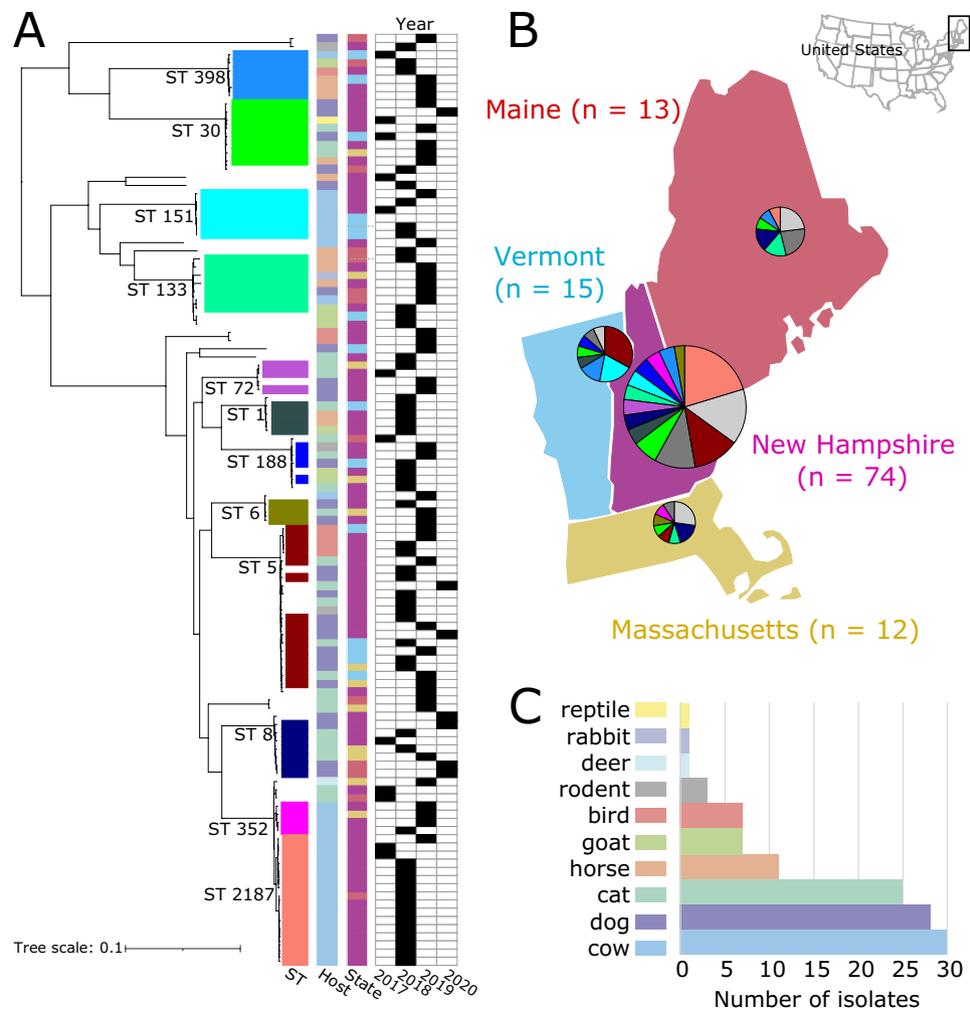


Figure 1. Phylogenetic diversity and sources of animal-associated *S. aureus* in New England. **(A)** Midpoint-rooted maximum likelihood phylogenetic tree. Tree scale represent the number of substitutions per site. For visual clarity, only those STs with at least three isolates are labeled. **(B)** Geographical distribution of STs. Pie charts show the proportion of major STs shown in panel A. Light gray includes STs consisting of a single isolate and novel STs, while dark gray represents STs consisting of two isolates. The number in parentheses indicates the number of isolates from each state. Maps were created using QGIS v3.22 (<https://www.qgis.org/en/site/>) **(C)** Number of isolates from different animal hosts.

Resistance-related genes associated with antibiotic efflux systems were ubiquitous in our dataset. We identified 11 such acquired genes, of which seven were found in all 114 genomes. These seven genes included *norA* (fluoroquinolone resistance), *tet(38)* (tetracycline resistance), *arlRS* (fluoroquinolone resistance), *lmrS* (multi-drug resistance) and *mepAR* (multidrug and tigecycline resistance). The expression of *norA* is affected by the two-component regulatory system *ArlRS*²³. *lmrS* confers resistance to aminoglycosides, macrolides, phenicols, diaminopyrimidine and oxazolidinone²⁴. *mepA* is an efflux protein regulated by *mepR* and part of the *mepRAB* cluster, which makes up the multidrug and toxin extrusion (MATE) mechanism in *S. aureus*²⁵. Also widely detected was *mgrA* (also known as *norR*), which is a regulator for *norA* and *tet(38)*²⁶. The *mgrA* gene was detected in all genomes except two.

Other resistance genes were less frequently detected. The gene *fosB* (fosfomycin resistance) was detected in 60 genomes and was present in some genomes from STs 5, 6, 8, 30, 72 and 133. We detected the gene *blaZ* (beta-lactam resistance) in 43 genomes and was present in some genomes from STs 5, 6, 8, 30, 72 and 398. In all, every genome carried at least eight resistance-related genes, of which seven genomes have at least 13 resistance related genes (Fig. 2B). Among these less common genes, their distribution across disparate parts of the phylogeny indicates the potential for multiple independent HGT events across different genetic backgrounds, rather than the spread of resistance through clonal expansion.

We next compared the frequency of resistance-related genes among isolates from different animal species. To maintain consistency in our comparison, we focused only on those animals with the highest number of isolates (cats, dogs). We detected a median of ten resistance-related genes (range = 8–15) in *S. aureus* from cats, ten resistance-related genes (range = 8–18) in *S. aureus* from dogs and eight resistance-related genes (range = 8–11) in

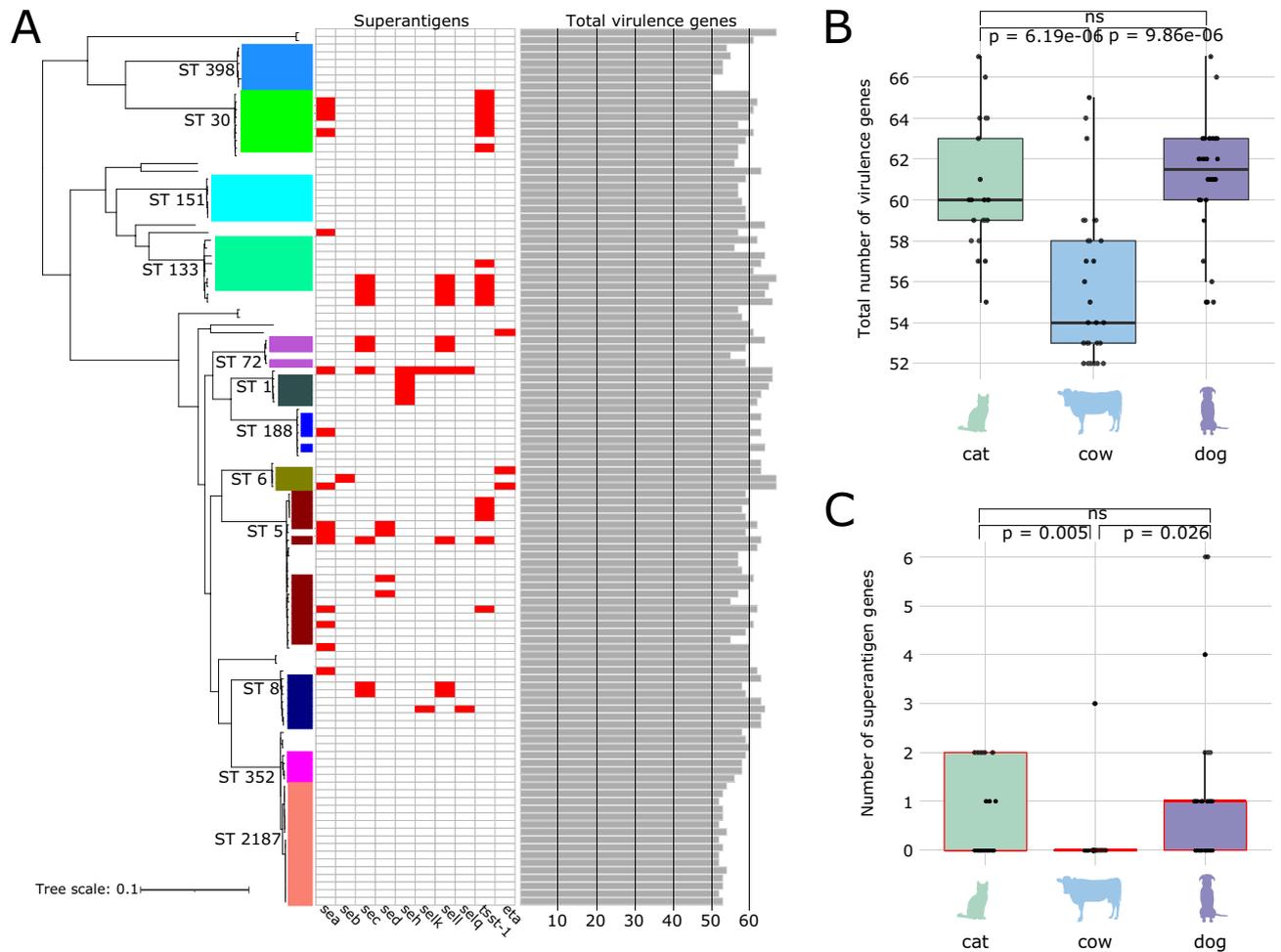


Figure 3. Virulence profiles of animal-associated *S. aureus* in New England. **(A)** Phylogenetic distribution of superantigen genes (shown in presence [red] or absence [white] matrix) and total number of virulence genes (shown in bar plots). **(B)** Number of virulence genes per *S. aureus* genome in cats, cows and dogs. **(C)** Number of superantigen genes per *S. aureus* genome in cats, cows and dogs. Significance was estimated using Welch's t-test.

diverse STs or lineages suggest multiple independent HGT events. The most common superantigen genes were *tsst-1* ($n = 17$ genomes in STs 5, 30 and 133), *sea* ($n = 15$ genomes in STs 5, 6, 30 and 188) and *sec* ($n = 10$ genomes in STs 5, 8, 72 and 133). STs 151, 352, and 2187 which were isolated exclusively from cows contained no superantigen genes, while ST 2187 exhibited fewer virulence genes and was additionally void of superantigen genes. A notable virulence factor in *S. aureus* is the Pantone-Valentine Leukocidin (PVL) pore-forming cytotoxin assembled by the genes *lukF-PV* and *lukS-PV*²⁹. PVL toxin-producing strains causes leukocytolysis and tissue necrosis and are often associated with community acquired MRSA infections in humans³⁰. In our dataset, we detected 96 and 3 genomes that carry *lukF-PV* and *lukS-PV*, respectively.

We also compared the number of virulence genes among isolates from the three most common animal hosts. Staphylococcal virulence genes were unevenly distributed among the three animal hosts. Isolates from cats harbored a median of 60 virulence genes (range = 55–67), 62 (range = 55–67) in dogs and 54 (range = 52–65) in cows (Fig. 3B). The total number of virulence genes varied between strains from cats and cows, as well as between isolates dogs and cows ($p = 6.19e-06$ and $9.86e-06$ respectively; Welch's t-test), but not between isolates from dogs and cats ($p = 0.331$; Welch's t-test) (Fig. 3B). When superantigens were considered, the number of superantigen genes carried by a genome ranged from 0–2, 0–3 and 0–6 in isolates from cats, cows and dogs, respectively (Fig. 3C). There was also no significant difference in the number of superantigen genes between strains from dogs and cats ($p = 0.366$; Welch's t-test). Three superantigen genes (*sec*, *sell*, *tsst-1*) were found in isolates from dogs, cats and cows. Six superantigen genes (*eta*, *sea*, *sed*, *seh*, *selk*, *selq*) were found only in dogs and cats, but not cows. Lastly, *seb* was found only in a single dog.

Widespread gene sharing between animal-associated *S. aureus*. Evidence for HGT in *S. aureus* and its contributions to adaptation in animal hosts has been demonstrated previously^{16,31,32}. However, little is known about the spread of potentially transferrable resistance and superantigen genes among the animal hosts of *S. aureus*. We mapped the presence or absence of these genes shared between isolates from the three most

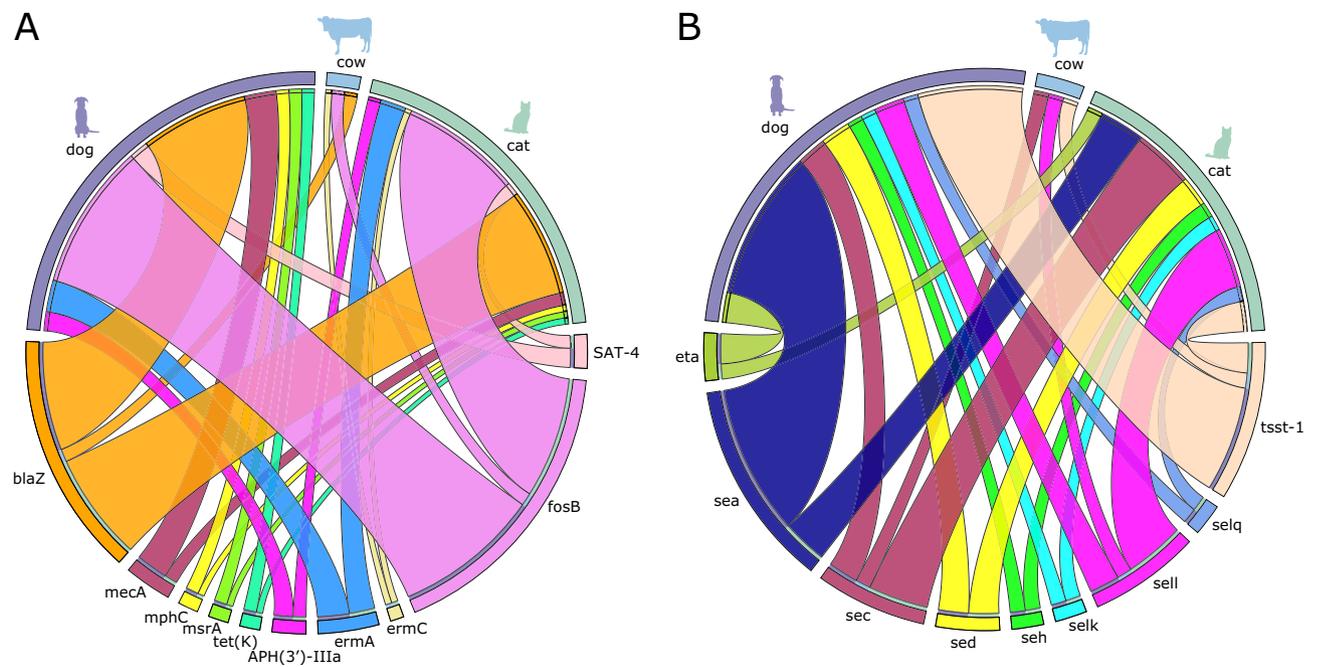


Figure 4. Distribution of antibiotic resistance (A) and superantigen (B) genes in *S. aureus* from cats, cows, and dogs. Only genes present in more than one isolate are shown and only the three most common animals were included. The outer ring of the upper half of circos plots represent the number of isolates sampled from dogs (purple), cows (light blue), and cats (light green). The outer ring of the bottom half of the circos plots represent the number of isolates that carry the shared antibiotic resistance genes (panel A) and superantigen genes (panel B). Connecting lines between the specific gene and the animal host are shown if the gene was detected in isolates from any of the three animal hosts.

common animal hosts (cats, cows, dogs). For the 15 resistance genes found in more than one isolate, the genes *blaZ* and *fosB* were present in isolates from all three animal hosts (Fig. 4A). We detected *blaZ* in 2, 17 and 17 isolates from cows, cats and dogs, respectively. We detected *fosB* in 2, 19 and 23 isolates from cows, cats and dogs, respectively. The genes which were detected in isolates from both cats and dogs, but not cows, were APH(3')-IIIa, *ermA*, SAT-4, *mecA*, *mphC*, *msrA*, and *tet(K)*. The gene *ermC* was present in isolates from cats and cows, but not dogs. Lastly, there were no genes that were shared between dogs and cows, but not cats.

Staphylococcal superantigen genes are often located in mobile genetic elements, such as prophages, transposons, plasmids and pathogenicity islands^{33,34}, thus facilitating their mobilization. We found that superantigen genes were widely distributed among isolates from different animal hosts (Fig. 4B). We detected *sec* in 1, 4, and 2 isolates from cows, cats and dogs, respectively. We detected *tsst-1* in 1, 2, and 7 isolates from cows, cats and dogs, respectively. The genes *sea*, *sed*, *seh*, *selk*, *selg*, and *eta* were present in isolates from both dogs and cats, but not cows. No superantigen genes were shared between isolates from dogs and cows, but not cats. Lastly, no superantigen genes were shared between isolates from cats and cows, but not dogs. Overall, these results showed widespread gene sharing between *S. aureus* gene pools from different host species.

Discussion

While humans are considered its primary reservoir, *S. aureus* can readily cross species barriers and infect new hosts¹⁶. Host-jumping events of bacterial pathogens are likely amplified by agricultural intensification, habitat encroachment and animal domestication^{35,36}. In the case of *S. aureus*, the range of eukaryotic species that it can colonize as a commensal or opportunistic pathogen remains unclear. Regardless, the remarkable capacity of *S. aureus* to adapt to new or multiple hosts thus makes it a formidable bacterium that can threaten animal health, agriculture and the economy. Host-jumping events are often associated with acquisition of genetic elements from host-specific gene pools that confer traits required for survival and adaptation in the new host niche¹⁶. These traits include a variety of virulence factors such as superantigens that can be used to manipulate innate and adaptive immune responses²⁸. For example, HGT in *S. aureus* mediated by mobile genetic elements occurs rapidly after a host-jumping event, potentially affecting the innate immune response of the new host¹⁶. During animal colonization, frequent HGT is facilitated by few genetic barriers to HGT in vivo (e.g., restriction–modification systems) and the successful replication and integration of different mobile elements^{37,38}.

Here, we sought to elucidate the population genomic structure of 114 *S. aureus* isolates sampled from diseased animals in New England, USA from 2017–2019. We found that many multidrug resistant STs were detected in multiple wild and domestic animals. Notable were STs 5, 8 and 30, which are major *S. aureus* clones that are implicated in nosocomial and community-associated infections in humans^{3,39}. ST133 isolates has been identified in healthy donkeys destined for food consumption in Tunisia⁴⁰, caprine and ovine animals in Australia⁴¹, and the

gut of healthy humans in Spain⁴². ST2187 has a long history of association with cows, and more specifically their milk^{43,44}. Although some *S. aureus* lineages are specifically adapted to a narrow host range on a short evolutionary time scale⁴⁵, as in the case of STs 151, 352 and 2187 in our study, this may be due to uneven and scarce sampling done in animals, particularly of wildlife species. Also notable is that 43 out of 114 genomes from six STs presented *blaZ*. For comparison, high penicillin susceptibility rate has been reported in MSSA from bloodstream infections in humans, primarily from CCs 5 and 398⁴⁶. Penicillin may be considered a therapeutic option in the treatment of animal infections, but ST identity should be carefully considered as *blaZ* appears to be more phylogenetically widespread in animal-associated *S. aureus*. As for the two isolates whose genomes did not contain the *mecA* gene but were phenotypically tested as MRSA, a similar finding has been reported in four isolates from the Scottish MRSA Reference Laboratory and which has been posited to indicate the existence of alternative mechanisms of beta-lactam resistance⁴⁷. Whole genome sequencing also revealed that those isolates phenotypically tested as MSSA often harbor numerous antibiotic resistance determinants. This result highlights the need to further investigate resistance characteristics beyond methicillin resistance in animal-associated *S. aureus*, which are often overlooked in many surveillance studies. In addition, such efforts will be instrumental in advancing the One Health concept, focused on the interconnectedness of animal, human and environmental well-being¹⁰.

Genes that encode for antibiotic resistance and superantigens were shared not only between divergent genetic backgrounds (or STs) of *S. aureus* but also between the animal hosts in which they reside. Hence, the ability of *S. aureus* to colonize multiple animal hosts means that mobilizable DNA may be disseminated more widely, creating a shared pool of resistance and virulence that is not limited by the animal hosts that harbor them. Our study also showed that cats, cows and dogs were frequent carriers of numerous *S. aureus* clones, each with distinct repertoire of resistance and superantigen genes. Hence, these animals are a major reservoir of clinically relevant genes and high-risk clones that can be transmitted to humans through frequent close contact. Overlapping ecological niches and/or physical proximity between animal hosts (e.g., pets in the same household, livestock animals in the same or nearby farms, interactions at the interface between wild and domestic animals) can certainly promote multi-host colonization and frequent gene sharing between *S. aureus* isolates. Domestic animals can therefore act as melting pots whereby genetic elements from various *S. aureus* lineages are combined in new genetic backgrounds. A variety of genetic assortments can promote the rapid emergence of high-risk clones with novel phenotypic characteristics. Large-scale genomic changes derived through HGT can generate “hopeful monsters” that may potentially cause public and animal health threats in ways that are hard to predict^{48,49}. Because *S. aureus* can also acquire DNA from other *Staphylococcus* species with which it shares its niche with^{31,50}, its gene pool may be further augmented with DNA that can confer additional adaptive or pre-adaptive features.

There are limitations in our study that need to be acknowledged. We recognize the sampling bias in our dataset that heavily favored isolates from cats, cows and dogs. Moreover, most isolates in our dataset were collected in New Hampshire where NHVDL is located, resulting in comparatively low sample sizes from the surrounding regions. Such bias did not allow us to carry out a more systematic analyses of the host distribution of STs and patterns of gene sharing between animal hosts that included other eukaryotes. The limited number of eukaryotic hosts means that those STs identified as host-restricted may in fact be found in multiple animal species. This also means that certain STs were overrepresented and rare ones were overlooked. Wildlife-associated *S. aureus* may likely harbor novel genetic variants or mobile genetic elements that pose an unknown level of risk to humans, companion animals and livestock. The structure and gene content of mobile genetic elements, such as pathogenicity islands and phages, and how they shape patterns of gene sharing in animal-associated *S. aureus* should also be investigated. Future work should therefore include a broader surveillance of *S. aureus* in other less commonly studied domestic animals and wildlife species, especially those species that often interface with livestock and/or exist at the junction of urban and natural landscapes. It is likely that the known range of species that *S. aureus* can colonize will expand as studies continue to examine *S. aureus* in wild animals. Our dataset also included only isolates from disease cases; hence, we do not have data to describe the extent of *S. aureus* carriage in animals and its contributions to the overall population genetic structure. Comparison of *S. aureus* in carriage and infections is critical to understanding the genetic basis of pathogenicity and hence, reduce the threats to animal health. Lastly, our study encompassed only three years of bacterial sampling, which was not sufficient to elucidate the long-term evolution of highly virulent and/or multidrug resistant lineages. Future investigations therefore necessitate close monitoring of high-risk clones and the underlying reasons that contribute to their persistence or expansion in the population.

In summary, our findings highlight the role of animals in disseminating resistance and virulence determinants of *S. aureus*. The remarkable ability of *S. aureus* as a versatile, multi-host pathogen lies partly on its ability to acquire and disseminate genetic material between lineages and between animal hosts within a short period of time. This study reveals widespread gene sharing between bacterial strains colonizing different animal hosts and highlights the need for routine surveillance to capture the dynamic genetic context of *S. aureus*.

Methods

Sample collection in New England. The New England *S. aureus* collection consisted of 133 isolates that were retrospectively sampled from September 2017 through March 2020. Isolates were obtained as culture swabs from routine clinical specimen submissions to the New Hampshire Veterinary Diagnostic Laboratory (NHVDL), New Hampshire, USA. The clinical specimens were received from multiple veterinary practices from the states of Connecticut, New Hampshire, Maine, Massachusetts and Vermont. These states are located in the northeastern part of the country. All isolates were from animals with confirmed clinical infections. No live vertebrates were used in this study; hence, the NHVDL was exempt from the IACUC approval process. Pure isolates were cultured in commercially prepared tryptic soy agar with 10% sheep red blood cells and brain heart infusion broth. Initial species identification was carried out using matrix-assisted laser desorption/ionization time-of-

flight mass spectrometry (MALDI-TOF MS) using the Bruker Biotyper instrument and following manufacturer's protocols. Species assignments were made by comparing mass spectra of our samples to two libraries of reference spectra RUO library 6903(V6) and 7311(V7) available in the Bruker MBT Compass (Bruker Daltonics, Bremen, Germany). The most common sites of isolation included skin, ears, nasal passages, milk and wounds. Associated metadata information for each isolate are included in Supplementary Table S1. All isolates were stored in DMSO solution in -80°C .

Methicillin susceptibility screening. In vitro screening for ceftioxin and oxacillin resistance was carried out using the Kirby Bauer disc diffusion technique. We followed the breakpoint guidelines of the most current Clinical and Laboratory Standards Institute using ceftioxin, which is used as the official predictor of methicillin resistance for *S. aureus*⁵¹. For isolates identified as methicillin resistant, we determined the presence of the penicillin-binding protein PBP2 using a commercial latex agglutination test (MASTALEX MRSA Latex Kit, MAST, UK) following manufacturer's guidelines. We confirmed the presence of the *mecA* gene by screening the genome sequence of each isolate (described below).

DNA extraction and whole genome sequencing. DNA extraction was carried out using the Zymo Research Quick-DNA Fungal/Bacterial Miniprep Kit (Irvine, California) following the manufacturer's protocol. We quantified DNA concentration using a Qubit fluorometer (Invitrogen, Grand Island, NY). We prepared DNA libraries using the Nextera XT protocol with 1 ng of genomic DNA per isolate. Samples were sequenced using the Illumina HiSeq platform (San Diego, California) to produce 250 bp paired end reads. Sequencing was carried out at the Hubbard Center for Genome Studies at the University of New Hampshire.

Genome assembly and annotation. We assembled all genomes using Shovill v1.1.0 (<https://github.com/tseemann/shovill>). Shovill is a series of methods that includes subsampling read depth down to 150X, trimming adapters, correcting sequencing errors and assembling using SPAdes v3.13.0⁵². We used QUAST v5.0.2⁵³ and CheckM v1.1.3⁵⁴ to assess the quality of our sequences and exclude genomes with $<90\%$ completeness and $>5\%$ contamination. We also excluded assemblies with >200 contigs and an $N50 < 40,000$ bp. Only 19 out of 133 genomes did not pass these quality thresholds and were therefore excluded from all downstream analyses. The remaining 114 genomes were annotated using Prokka v1.14.6⁵⁵. Genomes were compared to the *S. aureus* reference genome (NCBI Reference Sequence: Accession No. NC_007795.1) with the program FastANI v1.32 to confirm species identification using the $>95\%$ average nucleotide identity (ANI) threshold⁵⁶.

Pan-genome analysis and phylogenetic reconstruction. We used Roary v3.13.0⁵⁷ to characterize core genes and accessory genes that make up the pan-genome⁵⁷. To balance the tradeoff between inferring robust phylogenetic relationships versus accounting for assembly errors, we included core genes if they were present in $\geq 99\%$ of the genomes. Nucleotide sequences of each orthologous gene family were aligned using MAFFT v7.475⁵⁸. Aligned core genes were concatenated to generate a core genome alignment. Phylogenetically informative single nucleotide polymorphisms (SNPs) in the core genome alignment were extracted using SNP-sites⁵⁹. We used the core SNP alignment to construct a maximum likelihood phylogenetic tree using RAxML v8.2.12⁶⁰ employing a general time-reversible nucleotide substitution model⁶¹ and four gamma categories for rate heterogeneity. Phylogenetic trees were visualized using the online platform Interactive Tree of Life (IToL)⁶².

In silico sequence typing, detection of resistance genes, virulence genes and SCCmec. Using the contig files, we determined the multilocus sequence types (ST) for all genomes used in this study using the program mlst v2.19.0 (<https://github.com/tseemann/mlst>). STs pertain to allelic profiles that characterize nucleotide differences in partial sequences of seven housekeeping genes⁶³. In *S. aureus*, these seven genes consist of *arcC*, *aroE*, *glpF*, *gmk*, *pta*, *tpi* and *yqjL*⁶⁴. Allelic profiles of the genomes used in this study were compared to those in the *S. aureus* MLST database (<https://pubmlst.org/>)⁶⁵. We screened for the presence of horizontally acquired antibiotic resistance genes and virulence factors using ABRicate v1.0.1 (<https://github.com/tseemann/abricate>) utilizing the Comprehensive Antibiotic Resistance Database (CARD)¹⁷ and the Virulence Factor Database (VFDB)⁶⁶. Finally, we used SCCmecFinder⁶⁷ implemented in SCCion v0.1 (<https://github.com/estei/nig/sccion>) to determine the presence and type of the mobile genetic element SCCmec. We used the minimum thresholds of $>60\%$ for sequence coverage and $>90\%$ sequence identity to identify the SCCmec. Visualization of the distribution of resistance and virulence genes was carried out using Circos⁶⁸.

We used the default parameters for each program unless indicated otherwise.

Statistical tests. We used Welch's t-test to compare the number of resistance and virulence genes of isolates from different animal hosts. Results were considered significant when $p < 0.05$.

Data availability

Raw reads of the 114 *S. aureus* genomes have been deposited in the Sequence Read Archive (SRA) in the National Center for Biotechnology Information (NCBI) under BioProject accession No. PRJNA741582. Sample accession numbers and associated metadata are listed in Supplementary Table S1.

Received: 19 November 2021; Accepted: 4 March 2022

Published online: 15 March 2022

References

- Klein, E. Y., Sun, L., Smith, D. L. & Laxminarayan, R. The changing epidemiology of methicillin-resistant *Staphylococcus aureus* in the United States: a national observational study. *Am. J. Epidemiol.* **177**, 666–674 (2013).
- Tong, S. Y. C., Davis, J. S., Eichenberger, E., Holland, T. L. & Fowler, V. G. *Staphylococcus aureus* infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin. Microbiol. Rev.* **28**, 603–661 (2015).
- Turner, N. A. *et al.* Methicillin-resistant *Staphylococcus aureus*: an overview of basic and clinical research. *Nat. Rev. Microbiol.* **17**, 203–218 (2019).
- Fitzgerald, J. R. Livestock-associated *Staphylococcus aureus*: origin, evolution and public health threat. *Trends Microbiol.* **20**, 192–198 (2012).
- Faires, M. C., Traverso, M., Tater, K. C., Pearl, D. L. & Weese, J. S. Methicillin-resistant and -susceptible *Staphylococcus aureus* infections in dogs. *Emerg. Infect. Dis.* **16**, 69–75 (2010).
- Loncaric, I. *et al.* Increased genetic diversity of methicillin-resistant *Staphylococcus aureus* (MRSA) isolated from companion animals. *Vet. Microbiol.* **235**, 118–126 (2019).
- Monecke, S. *et al.* Diversity of *Staphylococcus aureus* isolates in European wildlife. *PLoS ONE* **11**, e0168433 (2016).
- Heaton, C. J., Gerbig, G. R., Sensius, L. D., Patel, V. & Smith, T. C. *Staphylococcus aureus* epidemiology in wildlife: a systematic review. *Antibiotics (Basel)* **9**, 89 (2020).
- Haag, A. F., Fitzgerald, J. R. & Penadés, J. R. *Staphylococcus aureus* in animals. *Microbiol. Spectr.* **7**, 11 (2019).
- Zinsstag, J., Schelling, E., Waltner-Toews, D. & Tanner, M. From 'one medicine' to 'one health' and systemic approaches to health and well-being. *Prev. Vet. Med.* **101**, 148–156 (2011).
- Spoor, L. E. *et al.* Livestock origin for a human pandemic clone of community-associated methicillin-resistant *Staphylococcus aureus*. *MBio* **4**, e00356-13 (2013).
- Lowder, B. V. *et al.* Recent human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*. *Proc Natl Acad Sci U S A* **106**, 19545–19550 (2009).
- Price, L. B. *et al.* *Staphylococcus aureus* CC398: host adaptation and emergence of methicillin resistance in livestock. *MBio* **3**, e0030511 (2012).
- Sieber, R. N. *et al.* Genome investigations show host adaptation and transmission of LA-MRSA CC398 from pigs into Danish healthcare institutions. *Sci Rep* **9**, 18655 (2019).
- Senghore, M. *et al.* Transmission of *Staphylococcus aureus* from humans to green monkeys in the Gambia as revealed by whole-genome sequencing. *Appl. Environ. Microbiol.* **82**, 5910–5917 (2016).
- Richardson, E. J. *et al.* Gene exchange drives the ecological success of a multi-host bacterial pathogen. *Nat. Ecol. Evol.* **2**, 1468–1478 (2018).
- Jia, B. *et al.* CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucl. Acids Res.* **45**, D566–D573 (2017).
- Katayama, Y., Ito, T. & Hiramatsu, K. A new class of genetic element, *staphylococcus cassette chromosome mec*, encodes methicillin resistance in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **44**, 1549–1555 (2000).
- International Working Group on the Classification of Staphylococcal Cassette Chromosome Elements (IWG-SCC). Classification of staphylococcal cassette chromosome *mec* (SCC*mec*): guidelines for reporting novel SCC*mec* elements. *Antimicrob. Agents Chemother.* **53**, 4961–4967 (2009).
- Lakhundi, S. & Zhang, K. Methicillin-resistant *Staphylococcus aureus*: molecular characterization, evolution, and epidemiology. *Clin Microbiol Rev* **31**(4), e00020-18 (2018).
- Urushibara, N., Aung, M. S., Kawaguchiya, M. & Kobayashi, N. Novel staphylococcal cassette chromosome *mec* (SCC*mec*) type XIV (5A) and a truncated SCC*mec* element in SCC composite islands carrying *speG* in ST5 MRSA in Japan. *J. Antimicrob. Chemother.* **75**, 46–50 (2020).
- García-Álvarez, L. *et al.* Methicillin-resistant *Staphylococcus aureus* with a novel *mecA* homologue in human and bovine populations in the UK and Denmark: a descriptive study. *Lancet Infect. Dis.* **11**, 595–603 (2011).
- Fournier, B., Aras, R. & Hooper, D. C. Expression of the multidrug resistance transporter *NorA* from *Staphylococcus aureus* is modified by a two-component regulatory system. *J. Bacteriol.* **182**, 664–671 (2000).
- Floyd, J. L., Smith, K. P., Kumar, S. H., Floyd, J. T. & Varela, M. F. *LmrS* is a multidrug efflux pump of the major facilitator superfamily from *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **54**, 5406–5412 (2010).
- Kaatz, G. W., McAleese, F. & Seo, S. M. Multidrug resistance in *Staphylococcus aureus* due to overexpression of a novel multidrug and toxin extrusion (MATE) transport protein. *Antimicrob. Agents Chemother.* **49**, 1857–1864 (2005).
- Truong-Bolduc, Q. C., Dunman, P. M., Strahilevitz, J., Projan, S. J. & Hooper, D. C. *MgrA* is a multiple regulator of two new efflux pumps in *Staphylococcus aureus*. *J. Bacteriol.* **187**, 2395–2405 (2005).
- Xu, S. X. & McCormick, J. K. Staphylococcal superantigens in colonization and disease. *Front Cell Infect. Microbiol.* **2**, 52 (2012).
- Tuffs, S. W., Haeryfar, S. M. M. & McCormick, J. K. Manipulation of innate and adaptive immunity by staphylococcal superantigens. *Pathogens* **7**, 53 (2018).
- Melles, D. C. *et al.* Pantone-Valentine leukocidin genes in *Staphylococcus aureus*. *Emerg. Infect. Dis.* **12**, 1174–1175 (2006).
- Boyle-Vavra, S. & Daum, R. S. Community-acquired methicillin-resistant *Staphylococcus aureus*: the role of Pantone-Valentine leukocidin. *Lab Invest.* **87**, 3–9 (2007).
- Méric, G. *et al.* Ecological overlap and horizontal gene transfer in *Staphylococcus aureus* and *Staphylococcus epidermidis*. *Genome Biol. Evol.* **7**, 1313–1328 (2015).
- Murray, S. *et al.* Recombination-mediated host adaptation by avian *Staphylococcus aureus*. *Genome Biol. Evol.* **9**, 830–842 (2017).
- Novick, R. P. Mobile genetic elements and bacterial toxinoses: the superantigen-encoding pathogenicity islands of *Staphylococcus aureus*. *Plasmid* **49**, 93–105 (2003).
- Alibayov, B., Zdenkova, K., Sykorova, H. & Demnerova, K. Molecular analysis of *Staphylococcus aureus* pathogenicity islands (SaPI) and their superantigens combination of food samples. *J. Microbiol. Methods* **107**, 197–204 (2014).
- Jones, B. A. *et al.* Zoonosis emergence linked to agricultural intensification and environmental change. *Proc. Natl. Acad. Sci. U S A* **110**, 8399–8404 (2013).
- Mourkas, E. *et al.* Agricultural intensification and the evolution of host specialism in the enteric pathogen *Campylobacter jejuni*. *Proc. Natl. Acad. Sci. U S A* **117**, 11018–11028 (2020).
- Roberts, G. A. *et al.* Impact of target site distribution for Type I restriction enzymes on the evolution of methicillin-resistant *Staphylococcus aureus* (MRSA) populations. *Nucl. Acids Res.* **41**, 7472–7484 (2013).
- McCarthy, A. J. *et al.* Extensive horizontal gene transfer during *Staphylococcus aureus* co-colonization in vivo. *Genome Biol. Evol.* **6**, 2697–2708 (2014).
- Monecke, S. *et al.* A field guide to pandemic, epidemic and sporadic clones of methicillin-resistant *Staphylococcus aureus*. *PLoS ONE* **6**, e17936 (2011).
- Gharsa, H. *et al.* High diversity of genetic lineages and virulence genes in nasal *Staphylococcus aureus* isolates from donkeys destined to food consumption in Tunisia with predominance of the ruminant associated CC133 lineage. *BMC Vet. Res.* **8**, 203 (2012).
- McMillan, K., Moore, S. C., McAuley, C. M., Fegan, N. & Fox, E. M. Characterization of *Staphylococcus aureus* isolates from raw milk sources in Victoria Australia. *BMC Microbiol.* **16**, 169 (2016).

42. Benito, D., Lozano, C., Gómez-Sanz, E., Zarazaga, M. & Torres, C. Detection of methicillin-susceptible *Staphylococcus aureus* ST398 and ST133 strains in gut microbiota of healthy humans in Spain. *Microb. Ecol.* **66**, 105–111 (2013).
43. Naushad, S., Nobrega, D. B., Naqvi, S. A., Barkema, H. W. & De Buck, J. Genomic analysis of bovine *Staphylococcus aureus* isolates from milk to elucidate diversity and determine the distributions of antimicrobial and virulence genes and their association with mastitis. *mSystems* **5**, (2020).
44. Thomas, A. *et al.* Prevalence and distribution of multilocus sequence types of *Staphylococcus aureus* isolated from bulk tank milk and cows with mastitis in Pennsylvania. *PLoS ONE* **16**, e0248528 (2021).
45. Matuszewska, M., Murray, G. G. R., Harrison, E. M., Holmes, M. A. & Weinert, L. A. The evolutionary genomics of host specificity in *Staphylococcus aureus*. *Trends Microbiol.* **28**, 465–477 (2020).
46. Mama, O. M. *et al.* Penicillin susceptibility among invasive MSSA infections: a multicentre study in 16 Spanish hospitals. *J. Antimicrob. Chemother.* **76**, 2519–2527 (2021).
47. Ba, X. *et al.* Novel mutations in penicillin-binding protein genes in clinical *Staphylococcus aureus* isolates that are methicillin resistant on susceptibility testing, but lack the *mec* gene. *J. Antimicrob. Chemother.* **69**, 594–597 (2014).
48. Croucher, N. J. & Klugman, K. P. The emergence of bacterial 'hopeful monsters'. *MBio* **5**, e01550-e1514 (2014).
49. Coyle, N. M. *et al.* A hopeful sea-monster: a very large homologous recombination event impacting the core genome of the marine pathogen *Vibrio anguillarum*. *Front. Microbiol.* **11**, 1430 (2020).
50. John, J., George, S., Nori, S. R. C. & Nelson-Sathi, S. Phylogenomic analysis reveals the evolutionary route of resistant genes in *Staphylococcus aureus*. *Genome Biol. Evol.* **11**, 2917–2926 (2019).
51. *CLSI Performance standards for antimicrobial disk and dilution susceptibility tests for bacteria isolated from animals. VET01.* (Clinical and Laboratory Standards Institute, 2018).
52. Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput Biol* **19**, 455–477 (2012).
53. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).
54. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
55. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
56. Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).
57. Page, A. J. *et al.* Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693 (2015).
58. Katoh, K., Asimenos, G. & Toh, H. Multiple alignment of DNA sequences with MAFFT. *Methods Mol. Biol.* **537**, 39–64 (2009).
59. Page, A. J. *et al.* SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb. Genom.* **2**, e000056 (2016).
60. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
61. Tavaré, S. Some probabilistic and statistical problems in the analysis of DNA sequences. *Am. Math. Soc. Lect. Math. Life Sci.* **17**, 57–86 (1986).
62. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucl. Acids Res.* **47**, W256–W259 (2019).
63. Maiden, M. C. *et al.* Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. U S A* **95**, 3140–3145 (1998).
64. Enright, M. C., Day, N. P., Davies, C. E., Peacock, S. J. & Spratt, B. G. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J. Clin. Microbiol.* **38**, 1008–1015 (2000).
65. Jolley, K. A., Bray, J. E. & Maiden, M. C. J. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res.* **3**, 124 (2018).
66. Liu, B., Zheng, D., Jin, Q., Chen, L. & Yang, J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucl. Acids Res.* **47**, D687–D692 (2019).
67. Kaya, H. *et al.* SCCmecFinder, a web-based tool for typing of staphylococcal cassette chromosome *mec* in *Staphylococcus aureus* using whole-genome sequence data. *mSphere* **3**, e0061217 (2018).
68. Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).

Acknowledgements

The authors thank the UAlbany Research Technology Services where all bioinformatics analyses were carried out. We are grateful to the Hubbard Center for Genome Studies at the University of New Hampshire for sequencing support. We are also grateful to the staff of NHVDL for laboratory support.

Author contributions

C.P.A. and S.A.B. designed the work. S.A.B. performed all bioinformatics analyses. J.T.S., J.L.M., J.B., D.N., and R.G. performed all sampling, culturing, and DNA extractions. C.P.A. and S.A.B. wrote the manuscript with contributions from all authors. C.P.A. guided the work. All authors read and approved the final manuscript.

Funding

The study was supported by the National Institutes of Health (NIH) Award No. 1R35GM142924 and startup research funds from the College of Arts and Sciences, University at Albany, State University of New York to CPA. JTS was supported by the University of New Hampshire 2020 Graduate School Dissertation Year Fellowship. The funders had no role in study design, data collection and analysis, decision to publish and preparation of the manuscript and the findings do not necessarily reflect views and policies of the authors' institutions and funders.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-08230-z>.

Correspondence and requests for materials should be addressed to S.A.B. or C.P.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022