A Shared Population of Epidemic Methicillin-Resistant *Staphylococcus aureus* 15 Circulates in Humans and Companion Animals

Ewan M. Harrison,a Lucy A. Weinert,a Matthew T. G. Holden,b John J. Welch,c Katherine Wilson,a Fiona J. E. Morgan,a Simon R. Harris,a Anette Loeffler,a Amanda K. Boag,a Sharon J. Peacock,b,e Gavin K. Paterson,f Andrew S. Waller,g Julian Parkhill,a Mark A. Holmesa

Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom; Wellcome Trust Sanger Institute, Hinxton, United Kingdom; Department of Genetics, University of Cambridge, Cambridge, United Kingdom; Royal Veterinary College, University of London, North Mymms, Hertfordshire, England; Department of Clinical Medicine, University of Cambridge, Cambridge, United Kingdom; School of Biological, Biomedical and Environmental Sciences, University of Hull, Hull, United Kingdom; Centre for Preventive Medicine, Animal Health Trust, Kentford, Newmarket, Suffolk, United Kingdom

**ABSTRACT** Methicillin-resistant *Staphylococcus aureus* (MRSA) is a global human health problem causing infections in both hospitals and the community. Companion animals, such as cats, dogs, and horses, are also frequently colonized by MRSA and can become infected. We sequenced the genomes of 46 multilocus sequence type (ST) 22 MRSA isolates from cats and dogs in the United Kingdom and compared these to an extensive population framework of human isolates from the same lineage. Phylogenomic analyses showed that all companion animal isolates were interspersed throughout the epidemic MRSA-15 (EMRSA-15) pandemic clade and clustered with human isolates from the United Kingdom, with human isolates basal to those from companion animals, suggesting a human source for isolates infecting companion animals. A number of isolates from the same veterinary hospital clustered together, suggesting that as in human hospitals, EMRSA-15 isolates are readily transmitted in the veterinary hospital setting. Genome-wide association analysis did not identify any host-specific single nucleotide polymorphisms (SNPs) or virulence factors. However, isolates from companion animals were significantly less likely to harbor a plasmid encoding erythromycin resistance. When this plasmid was present in animal-associated isolates, it was more likely to contain mutations mediating resistance to clindamycin. This finding is consistent with the low levels of erythromycin and high levels of clindamycin used in veterinary medicine in the United Kingdom. This study furthers the “one health” view of infectious diseases that the pathogen pool of human and animal populations are intrinsically linked and provides evidence that antibiotic usage in animal medicine is shaping the population of a major human pathogen.

**IMPORTANCE** Methicillin-resistant *Staphylococcus aureus* (MRSA) is major problem in human medicine. Companion animals, such as cats, dogs, and horses, can also become colonized and infected by MRSA. Here, we demonstrate that a shared population of an important and globally disseminated lineage of MRSA can infect both humans and companion animals without undergoing host adaptation. This suggests that companion animals might act as a reservoir for human infections. We also show that the isolates from companion animals have differences in the presence of certain antibiotic resistance genes. This study furthers the “one health” view of infectious diseases by demonstrating that the pool of MRSA isolates in the human and animal populations are shared and highlights how different antibiotic usage patterns between human and veterinary medicine can shape the population of bacterial pathogens.

*S. aureus* is part of the natural microbiota of humans and animals but has the potential to cause a broad spectrum of infections. The emergence and spread of methicillin-resistant *S. aureus* (MRSA) in both hospital and community settings pose a major threat to global health. Since the initial description in 1961, MRSA has spread globally with a small number of specific clones, including multilocus sequence types 239 (ST239), ST22, and ST8, causing the majority of the burden of disease (1, 2). Since the late 1990s, the role that both livestock and companion animals play as reservoirs and vectors for transmission of MRSA has become clearer (3–5). This relationship seems to be bidirectional, with human origins for the ST5 poultry lineage (6) and clonal complex 398 (CC398) originally infecting humans, jumping to livestock, and then back to humans (7, 8). More recently a number of “multihost” MRSA lineages with a novel mecA gene homologue, named mecC, were identified, which are capable of colonizing and infecting a broad range of mammalian and avian species (9–11). Companion animals, such as cats, dogs, and horses, are also frequently colonized by MRSA and can become infected (11–13). In dogs, reported MRSA carriage rates range from 0.7% in Portugal (14), to 2.3 to 9% in the United Kingdom (15, 16), and up to 20% in an outbreak in Canada (17). The prev-
alence of MRSA colonization of cats appears to be lower than that for dogs, with reported carriage rates of 0 to 4% (18), including 1.4% in Portugal (14) and 1.48% in the United Kingdom (13). Risk factors for MRSA infection in companion animals include contact with human MRSA carriers, the number of courses of antimicrobials received, length of time spent in veterinary clinics, and use of surgical implants (12, 15). MRSA lineages isolated from companion animals generally match the dominant lineages found in the human populations in the same geographical area: ST22 (epidemic MRSA 15 [EMRSA-15]) in the United Kingdom (19), Germany (20, 21), Portugal (22), and ST59/ST239 in China (23, 24). Molecular epidemiology using pulsed-field gel electrophoresis (PFGE) and spa typing has found that human and companion isolates are indistinguishable, suggesting transmission between humans and companion animals (25–29). Furthermore, a small number of studies have identified companion animals as the likely source for human MRSA infections (3, 28, 30). In the United Kingdom, ST22 (EMRSA-15) makes up the bulk of hospital-acquired cases of MRSA (31). A recent detailed phylogenomic study of the ST22 lineage demonstrated that it is likely to have emerged in the United Kingdom in the 1980s and then spread globally, with its initial success probably driven by the acquisition of fluoroquinolone resistance at a time of increasing fluoroquinolone usage (32). A single-locus variant of ST22 (ST2371) with Pantón-Valentine leukocidin (PVL) was recently tracked using whole-genome sequencing as spreading from a hospital outbreak (32). A single-locus variant of ST22 with Pantón-Valentine leukocidin (PVL) was previously reported by Holden et al. (32). Single nucleotide polymorphisms (SNPs) in the core genome were then used to reconstruct the phylogenetic relationships between the isolates. The substitution rate across the core genome under a constant population size model was estimated as $1.47 \times 10^{-6}$ per nucleotide site per year (95% Bayesian credible interval, $1.34 \times 10^{-6}$ to $1.60 \times 10^{-6}$), and for an exponentially growing population, it was estimated as $1.48 \times 10^{-6}$ (95% Bayesian credible interval, $1.36 \times 10^{-6}$ to $1.60 \times 10^{-6}$), similar to that previously reported for ST22 (32). No statistically significant difference was seen between the substitution rates for companion animal and human isolates (Fig. 1). Analysis of the phylogeny revealed that all of the companion animal isolates belonged to the previously described epidemic MRSA 15 (EMRSA-15) pandemic clade (ST22-A2 in reference 32).

RESULTS

Phylogenetics of companion animal isolates in comparison to human isolates. To investigate the relationship between ST22 isolates from companion animals and humans, we sequenced the genomes of 46 ST22 isolates from companion animals (42 canine and 4 feline) isolated between August 2003 and August 2007 (see Table S1 in the supplemental material). The isolates were from two large veterinary hospitals (The Royal Veterinary College, Hertfordshire, United Kingdom [24 isolates] and The Animal Health Trust, Suffolk, United Kingdom [5 isolates]), and a number of smaller veterinary practices throughout the United Kingdom (17 isolates). The companion animal isolates were from infections similar to those associated with ST22 in humans, with the majority of isolates coming from wound infections (including surgical site infections) (21) or skin and soft tissue infection (SSTI) (7). The collection also included isolates from urine (4), cerebrospinal fluid (CSF) (2), and nasal wash or discharge (2), and one isolate each from a bloodstream, heart valve, and joint infection. Only one isolate was from nasal carriage, and for six isolates, the clinical source was unknown. An additional 22 human CC22 isolates sequenced as part of other studies were also included in the analysis (34). The sequences were mapped against the ST22 reference genome HO 5096 0412 and then combined with the ST22 isolates previously reported by Holden et al. (32). Single nucleotide polymorphisms (SNPs) in the core genome were then used to reconstruct the phylogenetic relationships between the isolates. The substitution rate across the core genome under a constant population size model was estimated as $1.47 \times 10^{-6}$ per nucleotide site per year (95% Bayesian credible interval, $1.34 \times 10^{-6}$ to $1.60 \times 10^{-6}$), and for an exponentially growing population, it was estimated as $1.48 \times 10^{-6}$ (95% Bayesian credible interval, $1.36 \times 10^{-6}$ to $1.60 \times 10^{-6}$), similar to that previously reported for ST22 (32). No statistically significant difference was seen between the substitution rates for companion animal and human isolates (Fig. 1). Analysis of the phylogeny revealed that all of the companion animal isolates belonged to the previously described epidemic MRSA 15 (EMRSA-15) pandemic clade (ST22-A2 in reference 32).
The companion animal isolates were broadly distributed throughout the EMRSA-15 pandemic clade and, with one or two exceptions, clustered together with human isolates in United Kingdom-specific clades (Fig. 2A and 3). In all cases, human isolates were basal to those from companion animals, indicating that the evolutionary origin of the companion animal isolates was likely to be human. One particularly large clade of isolates contained human, cat, and dog isolates from three different veterinary practices in London and the southeast of England (see isolates in and around clades 1, 3, and 5 in Fig. 2A and 3).

**Isolates from large veterinary hospitals cluster together.** In general, the isolates from the larger veterinary hospitals clustered
together on the phylogenetic tree, while those from individual veterinary practices were not clustered with other companion animal isolates (see isolates from the Royal Veterinary College [RVC]) (isolates labeled “Hertfordshire” in Fig. 2A) and from the Animal Health Trust (AHT) (isolates labeled “Suffolk” in Fig. 2A). This was confirmed by randomization tests for significant clustering on the tree (isolates from large hospitals, P = 0.011; isolates from individual veterinary practices, P = 0.929). The RVC is situated in the southeast of England just outside London, and the majority of the isolates from the RVC (16 of 24 isolates) were also part of the large clade (described above) containing both human and companion animal isolates from London and the southeast (see isolates in and around clades 1, 3, and 5 in Fig. 2A). The phylogeny suggests that the RVC isolates were drawn from a population of EMRSA-15 isolates that was circulating widely throughout the human population of London and the southeast. Further analysis, focusing on isolates from the RVC, showed that multiple distinct clades with two or more closely related isolates separated by a small number of SNPs were circulating in the RVC between November 2003 and August 2006 (clades 1 to 5 in Fig. 2A). In two cases, in clades 2 and 3, three closely related isolates from different animals were isolated within ~2 months of each other (Fig. 2A and B), while isolates from clades 1, 4, and 5 were isolated from animals over a longer time periods (~6 to 12 months) (Fig. 2B).

Within each clade, isolates from a range of infections were present, demonstrating the ability of highly related isolates to cause a broad spectrum of disease in companion animals, as is the case in humans. In addition to the clusters of related isolates, a number of individual phylogenetically distinct isolates were also associated with infections in the RVC throughout the same period (Fig. 2A and B). A similar picture was seen in isolates from the other veterinary hospital in the study, the AHT, where two clades of closely related isolates were identified (isolates GBR_Suffolk_2005/a and GBR_Suffolk_2007/b in Fig. 2A). Phylogenetic analysis also identified a case of chronic infection or reinfection. Two wound isolates from the same cat taken ~20 days apart (GBR_Glasgow_2006 and GBR_Glasgow_2006a in Fig. 2), differed by only 3 SNPs, indicating that they shared a recent common ancestor. The level of diversity in these isolates is well within the observed variation of isolates in known cases of EMRSA-15 transmission (35), suggesting either a chronic wound infection or re-infection from the same source (Fig. 2A). Unfortunately, there was no further clinical information available to investigate the epidemiology.

Comparison of virulence factors and antibiotic resistance genes between human and animal isolates. In order to investigate if there was a genetic basis for the ability to colonize and cause disease in different hosts, we performed comparative genomic analysis on the ST22 isolates from the companion animals and human ST22 isolates. The entire collection of companion animal isolates was genotypically MRSA and harbored the SCCmec type IVb element as seen in the human ST22 isolates (32). All of the companion animal isolates had both of the fluoroquinolone resistance mutations in Ser80Phe in GrIA and Ser84Leu in GyrA previously described in human ST22 isolates (32). We assessed if the animal and human isolates shared the same virulence factors and antibiotic resistance genes previously identified in the human isolates (see Table S1 in the supplemental material) (32). We identified that the companion animal isolates were statistically more likely to have lost both the ßSa3 phage (containing the human-specific immune evasion genes sak, chips, and scin), with 76.1% ßSa3 positive, cf. 90.2% (Fisher’s exact test, P = 0.022), and the plasmid borne erm(C) erythromycin resistance gene, with 37% erm(C) positive, cf. 62% (Fisher’s exact test, P = 0.002) (Fig. 3). These analyses treat each isolate as an independent observation and so neglect the evolutionary history of the isolates as revealed by phylogenetic analysis. To control for phylogenetic non-independence, we treated the host state (human versus companion animal) and the presence of the ßSa3 phage (present versus absent) as binary traits and tested whether these traits evolved across the phylogeny in a correlated fashion. The results showed that a model with correlated evolution was not preferred to a model in which the traits evolved independently (Bayes factor 1, Bayes-Traits [73]), suggesting that companion animal isolates were no less likely to lose the ßSa3 phage than would be expected by chance. However, when we applied the same test to erm(C), correlated evolution between host state and loss/gain of erm(C) had substantial support (Bayes factor 9), providing further evidence that loss of erm(C) was associated with isolates coming from companion animals rather than humans. This host-species-associated loss is clearly visible on the phylogenetic tree, with the clades of companion animal isolates having lost the ErmC-encoding plasmid, while closely related and basal human isolates retained this plasmid (Fig. 3). We also identified three feline isolates (GBR_Glasgow_2006a, GBR_Glasgow_2006b, and GBR_Essex_2006) with genetic rearrangements disrupting the regulatory region upstream of erm(C), rendering its expression constitutive and thereby conferring resistance to clindamycin (36). Resistance to clindamycin in all three isolates was confirmed by phenotypic testing. Interestingly, the clindamycin-resistant isolates included the two isolates from the same cat described previously (GBR_Glasgow_2006a and GBR_Glasgow_2006b) that only differ by 3 SNPs. Furthermore, these two isolates each had different genetic rearrangements in the leader peptide, suggesting that two separate events led to resistance to clindamycin in these closely related isolates. The chronologically first isolate from the cat (GBR_Glasgow_2006a) had an insertion (IS) element inserted in the leader peptide, 66 bp upstream of the erm(C) start codon, while the second isolate, isolated ~20 days later, had a 58-bp deletion of the leader peptide, 69 bp upstream of the erm(C) start codon, (GBR_Glasgow_2006a). The third isolate, also from a cat (GBR_Essex_2006), contained an insertion of an IS element in the leader peptide region, 67 bp upstream of the start codon of erm(C). In contrast, none of 22 human isolates from the United Kingdom included in this study had the same or functionally equivalent rearrangements, as had been reported previously by Holden et al. (32) for the rest of United Kingdom human ST22 isolates and as found also in a wider study of the United Kingdom population of ST22 isolates (S. Peacock, personal communication).

No evidence of adaption to companion animals. To further identify if the isolates infecting the companion animals had mutations that might be associated with host adaptation, we used a genome-wide association (GWAS) approach. This approach uses multivariate regression analysis to identify SNPs in the core genome that are significantly associated with host state (e.g., SNPs present or absent in companion animal versus human isolates). The analysis showed very little genetic discrimination between isolates from different hosts (see Fig. S2 in the supplemental material), suggesting that the ST22 isolates had not undergone exten-
sive adaptation to companion animals. Five SNPs (three nonsynonymous, two synonymous) showed a more marked correlation with host state (companion animals) (see Fig. S3 and Table S3 in the supplemental material). However, ancestral state reconstruction of these SNPs across the phylogeny identified that they had each arisen only once in a single clade with both companion animal isolates and human isolates (Fig. 1 and 3, isolates around clades 1, 3 and 4). In addition, they had not been lost in human isolates from this clade (Fig. 2A). As such, it is most parsimonious to assume that the population structure alone was responsible for this correlation. Therefore, using this approach we were unable to identify any homoplastic SNPs potentially involved in host adaptation to the companion animals. We also investigated pseudogenes present in the ST22 isolates in an attempt to identify any particular gene inactivations that might be associated with adaptation to companion animals. However, we were unable to find any pseudogenes that had arisen more than once in companion animal isolates when taking into account the population structure (data not shown).

DISCUSSION
This is the first high-resolution genomic study of companion animal MRSA isolates, confirming what has been previously inferred from epidemiological data and lower-resolution molecular techniques (multilocus sequence typing [MLST], spa, and PFGE typing) that humans and companion animals readily exchange and share the MRSA isolates from the same population (16, 21, 25). Beyond this, we identified geographical structuring of human and
Harrison et al.

companion animal isolates from London and southeast England, suggesting that the same lineage was being exchanged between humans and companion animals in the local population. The widespread distribution of companion animal isolates throughout the EMRSA-15 phylogeny demonstrates that in all likelihood most EMRSA-15 isolates are capable of infecting both humans and companion animals. This ability to infect different hosts has been suggested to be an “extended-spectrum genotype” (20). A similar ability to cause infections in multiple species with limited genomic variation between strains has also been reported for ST130 (37) and CC398 (7). Speculatively, this ability to colonize and infect multiple species might be a beneficial adaption that provides EMRSA-15 isolates with an alternative pathway for transmission between human populations and that has contributed to the success of EMRSA-15, which is now the dominant lineage in much of Europe (38). Additionally, ST8 (USA300), the dominant lineage in North America, has been identified in cattle in Switzerland (39), cats and dogs in France (40), and cats, dogs, and pigs in the United States (41) and has been shown to be able to bind porcine corneocytes and persist in the nasal cavity of pigs (42). This suggests that a broad host range might be a common feature of successful S. aureus lineages and be characteristic of their long-term evolutionary history. Alternatively this might simply be due to a founder effect and simply reflect the predominant clones circulating in a particular geographical area.

We identified that, in a number of cases, isolates from the same veterinary hospitals clustered together, and highly related isolates were present in the same veterinary hospital over extended periods of time (~6 to 12 months). This suggests that EMRSA-15 isolates were persisting and being readily transmitted within the veterinary hospitals, as has been observed in human hospitals, suggesting that the pandemic potential of this clone is greater than generally acknowledged (33, 43). This finding is also consistent with the fact that all of the companion animal isolates in this study were from the predominately hospital-acquired EMRSA-15 (ST22-A) sublineage identified by Holden et al. rather than the predominately community-acquired broader ST22 group (32). The transmissibility of a hospital-acquired MRSA clone in a veterinary hospital setting demonstrates that MRSA prevention practices used in human medicine, such as search and destroy or blanket decolonization, may also be appropriate in veterinary practice (44). Further studies using whole-genome sequencing to investigate temporally paired human and companion isolates from veterinary hospitals and the human community are warranted to further understand the exact transmission dynamics between humans and companion animals.

A number of studies have identified genetic changes associated with adaptation of S. aureus to new host species (6, 7), including the most recent study of CC97 isolates jumping from cattle into humans (45). A common theme among these studies is the acquisition or loss of the β-toxin-converting phage (ðβSa3), which encodes the modulators of the human innate immune response; staphylokinase (SAK), staphylococcal complement inhibitor (SCIN), and chemotaxis inhibitory protein of S. aureus (CHIPS). We found no significant difference in the presence or absence of the ðβSa3 phage between human and companion animal isolates when correcting for shared evolutionary history. This lack of difference in the presence of ðβSa3 might be expected. While both CHIPS and SCIN are human specific and have significantly reduced activity against canine serum and neutrophils, respectively, SAK has been demonstrated to enhance the activation of canine plasminogen activation in vitro, suggesting that it might play a role in canine infections (46–48). No further loss or gain of virulence factors was associated with the isolates infecting companion animals, nor were any particular mutations (SNPs) detected in the core genome by a GWAS approach. This leads us to the conclusion that the core EMRSA-15 genome, present in isolates throughout the EMRSA-15 phylogeny, is sufficient to confer isolates with an extended host spectrum. EMRSA-15 has also been isolated from horses in Germany, United Kingdom, and Ireland (49, 50), goats in Spain (51), and a broad range of animal species, including a rabbit, parrot, turtle, and bat in Germany (52). Further studies including these isolates are needed to investigate if the same is true for isolates from diverse animal species.

One particularly interesting finding was that the companion animal isolates were significantly less likely to have maintained the plasmid-carried erm(C) gene, suggesting that the selective pressures for maintenance of the erm(C) plasmid are less in companion animals or that erm(C) is selected against. Previously, Harris et al. (33) reported the loss of the erm(C) plasmid in the CC22 isolates from two patients and in 18 of 20 colonies sequenced from a single colonized individual, suggesting this might be relatively common occurrence in the absence of selection (33, 53), as has been reported for other plasmids (54). Antibiotic resistance itself can come with significant fitness costs that can exact selective pressure elsewhere in the genome (55–57). In the United Kingdom, the use of erythromycin is rare in small animal veterinary medicine. During the period when the isolates in this study were collected (2005 to 2007), no erythromycin authorized for use in companion animals was sold in the United Kingdom (58–60). Another study by Mateus et al. in 2007 of 11 United Kingdom veterinary practices found that erythromycin usage was very rare and made up only 0.16% and 0.08% of prescriptions in dogs and cats, respectively (61). The picture in human medicine is dramatically different: 6.04 million prescriptions for erythromycin were made by general practitioners (GPs) between April 2005 and March 2007 in England alone (62). Thus, it is likely that the lack of usage of erythromycin in veterinary medicine in the United Kingdom meant there was a lack of selective pressure for the maintenance of erm(C) in companion animal isolates.

We also identified that three of the feline isolates had separate genetic rearrangements in the erm(C) leader peptide that are associated with the generation of clindamycin resistance. Previously, Holden et al. (32) suggested that the lack of clindamycin resistance in United Kingdom isolates from humans in comparison to isolates from other countries, including Germany and Sweden, was due to much higher clindamycin usage in these countries than in the United Kingdom. While there is very little use of lincosamides (of which clindamycin is a member) in human medicine in the United Kingdom (63), clindamycin is used widely in veterinary practice in the United Kingdom (61). A total of 2,533 kg of clindamycin authorized for use in companion animals was sold between 2005 and 2007 (58–60). Although clindamycin resistance was only present in three companion animal isolates, the fact that we identified two highly related isolates with different kinds of genetic rearrangements (IS insertion and deletion) combined with the significantly higher usage of clindamycin in veterinary practice does suggest that there might be a greater selective pressure for clindamycin resistance in veterinary settings as opposed to human health care settings.
We conclude that the intermingling of human and companion animals isolates within the EMRSA-15 pandemic clade suggests that isolates were capable of moving between the two populations, providing further support for the notion that companion animals may act as a reservoir for human MRSA infection and vice versa (64). This, combined with the findings related to difference in antibiotic resistance between human and companion animal isolates, highlights the importance of a “one health” view of infectious diseases—that the health of both human and animal populations are intrinsically linked.

MATERIALS AND METHODS

Isolates. Isolates were clinical cases submitted to diagnostic laboratories at the Animal Health Trust and Royal Veterinary Hospital (some isolates referred via Compton Paddock Laboratories) between August 2003 and August 2007. The isolates were subjected to in-house multilocus sequence typing, and all ST22 isolates were selected to be included for sequencing.

Whole-genome sequencing. Overnight cultures were grown in tryptic soy broth (TSB) at 37°C with 200-rpm shaking. Genomic DNA was extracted from cultures by using the MasterPure Gram-positive DNA purification kit (Cambio, United Kingdom). Illumina library preparation was carried out as described by Quail et al. (65). Hi-seq sequencing was carried out following the manufacturer’s standard protocols (Illumina, Inc., United States). Nucleotide sequences of the isolates have been deposited in the Sequence Read Archive database in the European Nucleotide Archive (see Table S1 in the supplemental material).

Antibiotic resistance testing. Antimicrobial susceptibility testing was performed using disk susceptibility testing according to BSAC criteria (BSAC Methods for Antimicrobial Susceptibility Testing, version 11.1, May 2012; British Society for Antimicrobial Chemotherapy, Birmingham, United Kingdom). NCTC12493 and NCTC6571 were used, respectively, as control resistant and susceptible isolates for clindamycin.

Phylogenetics and comparative genomics. Fastq files for the isolates were mapped against the ST22 MRSA reference genome HO 5096 0412 (EMBL accession no. HE681097) using SMALT (http://www.sanger.ac.uk/resources/software/smalt/) in order to identify single nucleotide polymorphisms (SNPs), as previously described (32, 35). SNPs located in mobile genetic elements (MGEs) were identified and removed from the alignment to generate a core genome (regions of the chromosome not excluded when MGEs were removed) (see Table S2 in the supplemental material). A maximum likelihood tree was generated from core genome SNPs by using RAXML (66). Trees were visualized and annotated with Figtree (http://tree.bio.ed.ac.uk/software/figtree/) and Interactive Tree Of Life (67). Comparison of the MGE content and virulence factors of the isolates was assessed by BLAST analysis against Velvet de novo assemblies by using MGEs, virulence and antibiotic resistance genes previously reported to be present in the ST22 lineage (32, 68). Comparative genomics were carried out using Velvet de novo assemblies with contigs realigned against HO 5096 0412 by Mauve (69) and manually inspected with the Artemis comparison tool (70).

Randomization tests. To carry out randomization tests of phylogenetic clustering, we asked whether the mean evolutionary distance between members in a specified collection of isolates was significantly smaller than the mean distance between the same number of randomly chosen isolates. The reported P values were derived from a null distribution estimated from 107 random samples of isolates, and all evolutionary distances were taken from the phylogeny depicted in Fig. 2A.

Time scale of evolution. We used the program BEAST v 1.7.5 (71) to estimate a dated phylogeny. To scale the rate of evolution, we constrained the tips of the phylogeny to the date that they were sampled in decimal years. If the day of the month was unknown, the first of the month was given as the date, and if the day and month were unknown, 1 June was used (see Fig. S1 in the supplemental material). Branch rates were drawn from an uncorrelated log normal distribution with mean rate assigned a uniform prior on the range 0 to 1. We assumed an HKY model of sequence evolution (72) with a gamma distribution of rate variation among sites. To model the relative node ages, we assumed either a constant or exponential population size coalescent tree prior. All other priors were assigned the defaults as specified in BEAUti v 1.7.5. We ran two separate Monte Carlo Markov chains (MCMCs) for each tree prior, using the maximum likelihood tree as a starting topology, until convergence in all parameters was reached and the burn-in was <10% of the entire run (∼3 × 106 iterations). The maximum clade credibility dated phylogeny (see Fig. S1) was generated from subsamples of the combined runs with burn-in removed. Estimates of the first date that ST22 S. aureus was found in dogs and cats were taken from the earliest node leading to a dog or cat isolate, respectively.

Correlated evolution of virulence genes. To test for correlated evolution between host the isolate was cultured from (human or companion animal) and virulence gene presence, we used the program BayesTraits (73), and the posterior sample of trees from our BEAST analysis using a constant coalescent tree prior. BayesTraits uses a continuous-time Markov model to estimate transition rates between the presence and absence of a virulence gene and between human and nonhuman hosts. We allowed the transition rates to evolve in either a correlated fashion (where the rate of change in one trait depends on the state found in the other trait) or independently. Posterior distributions of parameters were estimated from 4 × 106 iterations of the MCMC with default priors. After discarding burn-in, the marginal likelihoods of the dependent and independent models were obtained using the harmonic mean estimator in Tracer v 1.5 (74) to yield Bayes factors for the competing models.

Regression analysis. We used discriminant analysis of principal components (DAPC) in order to identify putative single nucleotide polymorphisms (SNPs) with a role in host adaptation (75). DAPC identifies linear combinations of biallelic SNPs that best discriminate between different groups (in our case, human, dog, or cat host states). We retained only SNPs with a >1% frequency of the minor allele (a total of 897/6,979 SNPs) and transformed the data using the dudi.pca function in the R package ade4, keeping 10 principal components (PCs) for use in the DAPC analysis (76). In the DAPC analysis using the R package adegenet, we again retained 10 PCs representing 60% of the total genetic variation and kept one discriminant function. Visualization of variables that contribute to between-group discrimination identified five outliers (see Fig. S3 in the supplemental material). The SNPs are shown in Table S3 in the supplemental material. The ancestral state of the SNPs was reconstructed by mapping the locations of the SNPs back onto the phylogeny.

SUPPLEMENTAL MATERIAL


Figure S1, TIFF file, 6.1 MB.
Figure S2, TIFF file, 0.4 MB.
Figure S3, TIFF file, 0.6 MB.
Table S1, XLSX file, 0.1 MB.
Table S2, XLSX file, 0.1 MB.
Table S3, XLSX file, 0.1 MB.

ACKNOWLEDGMENTS

We thank the core sequencing and informatics team at the Wellcome Trust Sanger Institute for sequencing of the isolates described in this study. We thank Ana Mateus and Katharina Staerk for providing further information about antimicrobial usage in veterinary practice.

This work was supported by a Medical Research Council Partnership grant (G1001787/1) held between the Department of Veterinary Medicine, University of Cambridge (M.A.H.), the School of Clinical Medicine, University of Cambridge (S.J.P.), the Moredun Research Institute, and the Wellcome Trust Sanger Institute (J.P. and S.J.P). S.J.P. receives support from the NIHR Cambridge Biomedical Research Centre. M.T.G.H., S.R.H. and J.P. were funded by Wellcome Trust grant no. 098051.

E.M.H. designed and carried out bioinformatic analysis and wrote the manuscript. L.A.W., M.T.G.H., and I.J.W. carried out bioinformatic and
Harrison et al.

statistical analysis and contributed to and critically revised the manuscript. K.W. carried out bioinformatics analysis. F.J.E.M. carried out isolation of samples and contributed to genome sequencing. S.R.H. provided analytical tools and contributed to the manuscript. A.L.K.B., and A.S.W. collected isolates and associated metadata and contributed to the manuscript. S.I.P., G.K.P., and J.P. contributed to the analysis and critically revised the manuscript. M.A.H. coordinated the study and wrote the manuscript. All authors read and approved the final manuscript.

REFERENCES


