RESEARCH ARTICLE

Staphylococcus aureus CC398: Host Adaptation and Emergence of Methicillin Resistance in Livestock

Lance B. Price,[a] Marc Stegger,[b] Henrik Hasman,[c] Maliha Aziz,[a] Jesper Larsen,[b] Paal Skytt Andersen,[b] Talima Pearson,[d] Andrew E. Waters,[a] Jeffrey T. Foster,[d] James Schupp,[a] John Gillece,[a] Elizabeth Driebe,[a] Cindy M. Liu,[a,d] Burkhard Springer,[e] Irena Zdovc,[f] Antonio Battisti,[g] Alessia Franco,[g] Jacek Żmudzki,[h] Stefan Schwarz,[i] Patrick Butaye,[j,k] Eric Jouy,[l] Constanca Pomba,[m] M. Concepción Porrero,[n] Raymond Ruimy,[o] Tara C. Smith,[p] D. Ashley Robinson,[q] J. Scott Weese,[r] Carmen Sofia Arriola,[s] Fangyou Yu,[t] Frederic Laurent,[u] Paul Keim,[a,d] Robert Skov,[b] and Frank M. Aarestrup[c]

Translational Genomics Research Institute (TGen), Pathogen Genomics Division, Flagstaff, Arizona, USA[a]; Department of Microbiological Surveillance and Research, Statens Serum Institut, Copenhagen, Denmark[b]; National Food Institute, Technical University of Denmark, Lyngby, Denmark[c]; Center for Microbial Genetics and Genomics, Northern Arizona University, Flagstaff, Arizona, USA[d]; Österreichische Agentur für Gesundheit und Ernährungssicherheit GmbH, Bereich Humanmedizin, Institut für Medizinische Mikrobiologie und Hygiene, Graz, Austria[e]; Veterinary Faculty, Institute of Microbiology and Parasitology, University of Ljubljana, Ljubljana, Slovenia[f]; Istituto Zooprofilattico Sperimentale delle Regioni Lazio e Toscana, Rome, Italy[g]; Department of Swine Diseases, National Veterinary Research Institute, Pulawy, Poland[h]; Institute of Farm Animal Genetics, Friedrich-Loeffler-Institut, Neustadt-Mariensee, Germany[i]; Department of Pathology, Bacteriology, and Poultry Diseases, Faculty of Veterinary Medicine, University of Ghent, Merelbeke, Belgium[j]; Unit of General Bacteriology, Veterinary and Agrochemical Research Centre, Brussels, Belgium[k]; Anses, Ploufragan Plouzané Laboratory, Ploufragan, France[l]; Interdisciplinary Centre of Research in Animal Health, Faculty of Veterinary Medicine, Lisbon, Portugal[m]; Centro de Vigilancia Sanitaria Veterinaria, Universidad Complutense de Madrid, Madrid, Spain[n]; University Paris-Diderot and National Reference Center, Bacterial Resistance in the Commensal Flora, Hôpital Bichat-Claude Bernard, Paris, France[o]; Center for Emerging Infectious Diseases, University of Iowa, Coralville, Iowa, USA[p]; Department of Microbiology, University of Mississippi Medical Center, Jackson, Mississippi, USA[q]; Ontario Veterinary College, University of Guelph, Guelph, Ontario, Canada[r]; Department of Preventive Veterinary Medicine, San Marcos Major National University, Lima, Peru[s]; Department of Laboratory Medicine, First Affiliated Hospital of Wenzhou Medical College, Wenzhou, China[t]; and Laboratory of Bacteriology, Hôpital de la Croix Rousse, National Reference Center for Staphylococci, Lyon, France[u]

ABSTRACT Since its discovery in the early 2000s, methicillin-resistant *Staphylococcus aureus* (MRSA) clonal complex 398 (CC398) has become a rapidly emerging cause of human infections, most often associated with livestock exposure. We applied whole-genome sequence typing to characterize a diverse collection of CC398 isolates ($n = 89$), including MRSA and methicillin-susceptible *S. aureus* (MSSA) from animals and humans spanning 19 countries and four continents. We identified 4,238 single nucleotide polymorphisms (SNPs) among the 89 core genomes. Minimal homoplasy (consistency index = 0.9591) was detected among parsimony-informative SNPs, allowing for the generation of a highly accurate phylogenetic reconstruction of the CC398 clonal lineage. Phylogenetic analyses revealed that MSSA from humans formed the most ancestral clades. The most derived lineages were composed predominantly of livestock-associated MRSA possessing three different staphylococcal cassette chromosome *mec* element (SCC*mec*) types (IV, V, and VII-like) including nine subtypes. The human-associated isolates from the basal clades carried phages encoding human innate immune modulators that were largely missing among the livestock-associated isolates. Our results strongly suggest that livestock-associated MRSA CC398 originated in humans as MSSA. The lineage appears to have undergone a rapid radiation in conjunction with the jump from humans to livestock, where it subsequently acquired tetracycline and methicillin resistance. Further analyses are required to estimate the number of independent genetic events leading to the methicillin-resistant sublineages, but the diversity of SCC*mec* subtypes is suggestive of strong and diverse antimicrobial selection associated with food animal production.

IMPORTANCE Modern food animal production is characterized by densely concentrated animals and routine antibiotic use, which may facilitate the emergence of novel antibiotic-resistant zoonotic pathogens. Our findings strongly support the idea that livestock-associated MRSA CC398 originated as MSSA in humans. The jump of CC398 from humans to livestock was accompanied by the loss of phage-carried human virulence genes, which likely attenuated its zoonotic potential, but it was also accompanied by the acquisition of tetracycline and methicillin resistance. Our findings exemplify a bidirectional zoonotic exchange and underscore the potential public health risks of widespread antibiotic use in food animal production.

Address correspondence to Lance B. Price, lprice@tgen.org.

Methicillin-resistant *Staphylococcus aureus* (MRSA) has been described in animals since 1972 (1), but a new lineage, clonal complex 398 (CC398), has emerged among livestock and begun colonizing and infecting humans. Human MRSA infections have been categorized into three groups based on their putative sources: health care-associated MRSA, community-associated MRSA,

and health care-associated MRSA with community onset. A fourth category has recently been added to describe human MRSA cases associated with exposure to livestock (livestock-associated MRSA [LA-MRSA]). Human colonization with LA-MRSA multilocus sequence type 398 (ST398) was first recognized among swine farmers in France and The Netherlands in the early 2000s (2, 3). Since those early reports, ST398 and closely related STs within CC398 have been reported in diverse livestock hosts in many countries around the world (4–8). Human cases of MRSA CC398 have also been increasing rapidly and now account for up to 25% of the total MRSA cases in some parts of The Netherlands (9). Given its rapid emergence and trajectory of increasing importance in humans, the evolutionary history of MRSA CC398 has relevance for the epidemiology of MRSA and global health.

Methicillin-susceptible *S. aureus* (MSSA) CC398 is prevalent among pigs in Europe (10), but the evolution and global dispersal of this group have yet to be clarified. A microarray-based study revealed that the core genomes of 6 CC398 isolates were distinct from those of more than 2,000 *S. aureus* isolates from humans (11), and it has been generally presumed that pigs or other animals are the natural hosts of CC398.

Attempts to better characterize the evolution and epidemiology of CC398 have been hampered by the limited resolution of conventional *S. aureus* typing methods, including multilocus sequence typing (MLST) and *spa* sequence typing. While MLST defines CC398 and is useful for placing CC398 in the context of other *S. aureus* clonal complexes, it is of no use for characterizing variation within the group. *spa* typing has revealed geographic clustering among CC398 isolates in Europe (12, 13); however, the limited number of common *spa* types among CC398 isolates and the potential for homoplasy in the *spa* gene data restrict its phylogenetic utility (14).

Whole-genome sequencing provides a superior genetic fingerprint, which can be used for source tracking and evolutionary studies. Thus far, whole-genome sequencing has been used to study the hospital transmission and spatiotemporal spread of ST239 (15). In this study, we applied whole-genome sequence typing (WGST) to a diverse collection of 89 CC398 isolates to study the origins and evolution of *S. aureus* CC398.

RESULTS

We sequenced the genomes of 88 *S. aureus* CC398 isolates (see Data set S1 in the supplemental material). Among the isolates, we sequenced an average of 2,651,848 bases (standard deviation [SD] = 80,311) at ≥10× coverage. Genomes were sequenced at an average depth of 104.36× (SD = 35.7, using the 2,872,582-base SO385 chromosome as a reference).

Rooting the CC398 tree using ST36 as the outgroup revealed that a cluster of four isolates, characterized by the t899 *spa* type, was the first lineage to diverge from the other CC398 isolates (see Fig. S1 in the supplemental material). Much of the similarity between the t899 cluster and the ST36 isolate was mapped to the same ~123,000-bp region surrounding the *spa* gene. However, this region was incongruent with the phylogenetic signal generated by the rest of the chromosome in the t899 isolates. Comparative genomic analysis with other STs suggested that this region was acquired horizontally from an ST9 donor (see Fig. S2 in the supplemental material). When single nucleotide polymorphisms (SNPs) from this region were excluded from the phylogenetic analysis, the t899 lineage clustered with a more derived clade of

European isolates and a clade of human-associated MSSA isolates from France, French Guiana, and the United States was identified as the first divergent CC398 lineage (see Fig. S3 in the supplemental material). This lineage was used to root the final CC398 WGST tree (Fig. 1). With the ST36 genome removed and excluding the SNPs from the 123,000-bp putative horizontally transferred region, we identified 4,238 SNPs, including 1,102 parsimony-informative SNPs with a consistency index (CI) of 0.9591. Among the SNPs, 3,552 were from coding regions (1,071 synonymous and 2,481 nonsynonymous).

The phylogenetic tree presented in Fig. 1 is a highly accurate depiction of the evolutionary relationships among the 89 CC398 strains included in this study (including the reference). The lack of homoplasy among informative SNPs (CI = 0.9591) obviated the need for additional measures of robustness such as bootstrapping (16). The most ancestral lineage (clade I) was composed entirely of MSSA strains from humans in North America, South America, and Europe. In addition, with the exception of one isolate (P23-14_SD4.1), the strains that accounted for the most ancestral lineages within clade II (II group of interest [II-GOI]) were human-associated *S. aureus* from China (including two MRSA strains isolated from Danish adoptees from China). All but one of the livestock-associated strains belonged to clade IIa, which was derived from the human-associated lineages. Clade IIa was composed of several lineages whose evolutionary relationships could not be determined due to poor hierarchical resolution. The lack of resolution was likely due to a rapid radiation following introduction into livestock, as homoplasy was exceedingly rare among the parsimony-informative SNPs. The isolates within clades IIa1 and IIa2 consisted almost entirely of European isolates, while the 10 remaining lineages consisted almost entirely of North American isolates. The IIa1i lineage was dominated by Danish isolates, one-third of which were MSSA, while the IIa1ii lineage consisted largely of MRSA isolates from several European countries, except Denmark. Interestingly, 6 of the 10 smaller lineages included MSSA strains isolated from turkey meat from the United States (IIa-GOI). In this report, we use the term "human associated" for isolates belonging to clade I and clade II-GOI (*n* = 19) and use the term "livestock associated" for isolates belonging to subclade IIa (*n* = 70).

Fifteen different *spa* types were identified among the 89 CC398 isolates, including t011, t034, t108, t567, t571, t899, t1250, t1451, t1793, t2876, t3085, t3625, t5462, t5463, and t5719 (Fig. 1). The two most common *spa* types, t011 and t034, represented 67% of the isolates. While some *spa* types were more common within individual clusters (e.g., t571 was disproportionately common among human-associated isolates), *spa* types were inconsistent with the overall CC398 phylogeny (Fig. 1).

Sixty-one percent (30/49) of the CC398 MRSA isolates harbored staphylococcal cassette chromosome *mec* element (SCC-*mec*) subtype Vc (5C2&5) containing the cadmium-zinc resistance gene *czrC*. All of the SCC*mec* Vc (5C2&5) cassettes were present in LA-MRSA strains. Of note, the *czrC* gene was also found in two livestock-associated MSSA isolates. The remaining LA-MRSA isolates carried SCC*mec* (sub)types IVa (2B), IVa (2B&5), IVc (2B), Vb (5C2&5), V*, and V**; a novel VII-like SCC*mec* cassette; and a nontypeable (NT) SCC*mec* cassette. The *mecA* gene was detected in only two human-associated isolates; in both cases, the *mecA* gene was coded within SCC*mec* subtype Vb (5C2&5). The type V* and V** SCC*mec* cassettes contained structurally dif-
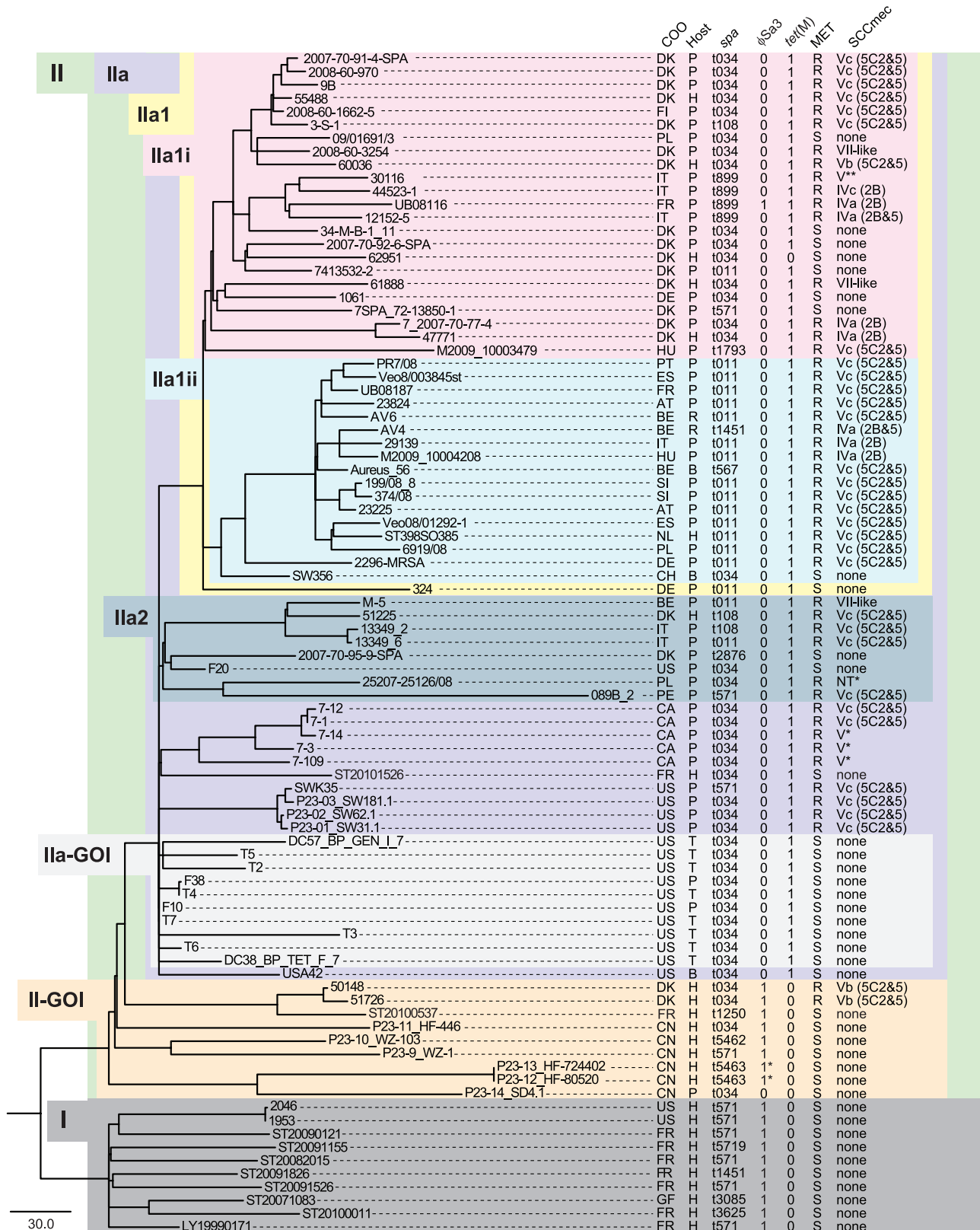
FIG 1 Maximum-parsimony tree of the 89 CC398 isolates (including ST398SO385) based on 4,238 total SNPs, including 1,102 parsimony-informative SNPs with a CI of 0.9591. Clades and groups of importance are labeled in a hierarchical fashion to facilitate description in the text. The tree was rooted with clade I based on an iterative selection process that identified this group as the most ancestral (see Materials and Methods). COO, country of origin; AT, Austria; BE, Belgium; CA, Canada; CH, Switzerland; CN, China; DE, Germany; DK, Denmark; ES, Spain; FI, Finland; FR, France; GF, French Guiana; HU, Hungary; IT, Italy; NL, The Netherlands; PE, Peru; PL, Poland; PT, Portugal; SI, Slovenia; US, United States; P, pig; H, human; R, horse; T, turkey; B, bovine; MET, methicillin susceptibility; R, resistant; S, susceptible.

ferent J1 regions that did not match the J1 regions associated with subtypes a to c. The type VII-like SCC*mec* cassette contained *ccr* type 5 (*ccrC*) and a class C1-like *mec* gene complex element previously identified in SCC*mec* type X (7C1-like) (17). Accordingly, this novel SCC*mec* type was referred to as VII-like (5C1-like) to distinguish it from archetypal SCC*mec* type VII (5C1). The NT cassette contained a class C1-like *mec* gene complex without any previously described *ccr* gene complex. The tetracycline resistance gene *tet*(M) was present in 99% (69/70) of the livestock-associated isolates but absent from the human-associated isolates (Fig. 1; see Data set S1 in the supplemental material).

The prophage integrase gene Sa3int was detected in 29 CC398 isolates. Phylogenetic analysis of the Sa3int sequences showed that they belong to three separate clusters; one clade was typical of φSa3 prophages, one was typical of φAvβ prophages, and the third was suggestive of a novel φSa3 integrase variant (see Fig. S4 in the supplemental material). φSa3 prophages in association with one or more human innate immunomodulatory genes were detected in 95% (18/19) of the human-associated *S. aureus* isolates (Fig. 1; see Data set 1 in the supplemental material). All 18 positive isolates were from human samples, whereas the single isolate lacking φSa3 prophages originated from a pig farm. All 10 isolates belonging to clade I carried *chp* and *scn* (type C φSa3 prophages), whereas 6 of 8 isolates belonging to clade II-GOI carried *sak*, *chp*, and *scn* (type B φSa3 prophages) and 2 isolates carried *scn* only (an IEC type not previously described). In comparison, only 1 of 70 isolates belonging to clade II harbored a φSa3 prophage in association with *sak*, *chp*, and *scn* (type B). Interestingly, 10 livestock-associated isolates belonging to IIa-GOI were largely from turkey meat samples and carried a φAvβ prophage along with the associated genes SAAV_2008 and SAAV_2009 but lacked human innate immunomodulatory genes carried by φSa3 prophages. Sa2int and the *lukF-lukS* genes carried by φSa2 prophages were present in 6 of 19 human-associated isolates. Conversely, all livestock-associated *S. aureus* isolates lacked the *lukF-lukS* genes.

DISCUSSION

Since its discovery, MRSA CC398 has been perceived as a livestock-associated pathogen; however, the WGST-based phylogeny presented here strongly suggests that the CC398 lineage originated in humans as MSSA and then spread to livestock, where it subsequently acquired the SCC*mec* cassette and methicillin resistance. The isolates that formed the most basal clades (I and II-GOI) on the WGST-based phylogenetic trees were almost all human-associated MSSA strains, suggesting that these isolates were the most ancestral of those tested in this study (Fig. 1). Likewise, the clade structure observed in the livestock-dominated IIa clade supports a rapid radiation as CC398 moved from humans to animals (see Fig. S5 in the supplemental material). Thus, livestock-associated CC398 infections in humans may be seen as a reintroduction to the original host.

Epidemiological data suggest that livestock-associated CC398 strains have lower transfer rates, and may be less virulent, in humans than other well-known STs (18). In this study, we showed that the *lukF-lukS* genes encoding Panton-Valentine leukocidin (PVL) were present in only 6 of the 89 genomes, all of which were human associated (see Data set S1 in the supplemental material). Strikingly, we found that all of the human-associated MSSA strains from clade I and clade II-GOI carried φSa3 in association with human innate immunomodulatory genes, whereas φSa3 was

identified in only one livestock-associated isolate (see Fig. S6 in the supplemental material). Instead, a φAvβ prophage and the associated genes SAAV_2008 and SAAV_2009 were identified among a group of mainly turkey meat isolates in the livestock-associated clade IIa-GOI. It therefore appears that φSa3 was lost prior to (or early in) the formation of clade II, while φAvβ was introduced into avian MSSA CC398 isolates thereafter (Fig. 1). The human innate immunomodulatory genes carried by φSa3 prophages play crucial roles in human niche adaptation (19, 20), whereas the φAvβ-carried SAAV_2008 and SAAV_2009 genes (encoding a putative ornithine cyclodeaminase and a putative membrane protease of the CAAX family, respectively) belong to the avian-niche-specific accessory gene pool for broiler chicken-associated *S. aureus* ST5 (21). The loss of human-niche-specific genes in livestock-associated isolates, including those from turkeys, may be a result of adaptation to nonhuman hosts. A similar natural history has been reconstructed for broiler chicken-associated *S. aureus* ST5, which appears to have been introduced from humans into the chicken-breeding system, transmitted vertically, and disseminated worldwide (21). The ST5 jump from humans to chickens also appears to have been followed by the acquisition of avian-niche-specific genes (including the SAAV_2008 and SAAV_2009 genes carried by φAvβ prophages) and partial loss of human-niche-specific genes (including human innate immunomodulatory genes carried by φSa3 prophages) (21).

The data presented here strongly suggest that CC398 acquired resistance to methicillin and tetracycline after the introduction to livestock from humans (see Fig. S7a and b in the supplemental material). The tetracycline resistance gene *tet*(M) was nearly universal among livestock-associated CC398 MRSA and MSSA isolates and completely missing from human-associated strains. Consequently, tetracycline use in food animal production is likely to select for livestock-associated *S. aureus* CC398 without differentially selecting for MRSA strains. MRSA can be selected for by a number of broad-spectrum cephalosporins that are used in food animal production in the United States and Europe. Likewise, zinc and other metals are frequently used in animal feed formulations and may coselect for MRSA CC398 strains that carry the *czrC* zinc resistance gene, as suggested previously (22). This hypothesis is supported by our findings that the vast majority of LA-MRSA strains carry SCC*mec* type Vc (5C2&5), which contains the *czrC* gene.

This study demonstrates the potential power of WGST for epidemiological investigations. For example, two of the Danish MRSA isolates came from infants adopted from China. Both isolates were *spa* type t034, which is consistent with the majority of Danish CC398 isolates from pigs and humans; however, WGST showed that the isolates shared a recent common ancestor with a French isolate and that this clade was derived from other clades within II-GOI that were strictly Chinese in origin (Fig. 1, II-GOI). Although the French isolate obscures these results, they are most consistent with a Chinese rather than Danish origin of the isolates. WGST revealed thousands of SNPs among the 89 CC398 strains. These mutations may provide robust phylogenetic signals for future epidemiological and epizootological investigations involving CC398 strains.

spa typing is routinely used for *S. aureus* epidemiology; however, in this study, homoplasy within the *spa* gene led to inconsistencies between the WGST CC398 phylogeny and *spa* types. Some

spa types, such as t571 and t034, were observed in distant clades of the highly accurate WGST phylogenetic tree (Fig. 1). The t899 isolates exemplified the limitations of any single-locus typing method, as the *spa* gene was part of a ~123,000-bp region of DNA acquired from a distantly related *S. aureus* clone. A similar observation was made previously with *S. aureus* ST239, which originated as a hybrid between ST8-like and ST30-like chromosomes (23). Here, reliance on *spa* typing would have incorrectly placed these isolates outside of CC398. Interestingly, the large horizontally acquired region observed among the t899 CC398 strains also carries the SCC*mec* cassette, thus possibly presenting an alternative mechanism for SCC*mec* dissemination among *S. aureus* strains.

In this study, we provide strong evidence that CC398 originated in humans as MSSA and then spread to livestock, where it acquired resistance to methicillin and tetracycline. Genomic analyses presented here, in conjunction with previous epidemiological data, suggest that the jump from humans to animals was followed by a decreased capacity for human colonization, transmission, and virulence, yet livestock-associated CC398 has been linked to an increase in MRSA infections in northern Europe. Further research is required to characterize the full scope of the genetic changes associated with the shift from humans to livestock. Likewise, additional research and surveillance are required to predict the public health impact of MRSA CC398 in the future.

MATERIALS AND METHODS

Bacterial isolates. This study included MRSA (*n* = 48) and MSSA (*n* = 40) CC398 isolates from 19 countries on four continents with strains from humans (*n* = 25) and livestock (*n* = 63, including strains from live animals, meat samples, and environmental contamination) (see Data set S1 in the supplemental material). A previously sequenced ST398 strain, SO385, from The Netherlands was used as the reference and included in all analyses (24).

MLST. MLST was performed as described previously (http://saureus.mlst.net/misc/info.asp) (25). STs were assigned through the MLST database (http://www.mlst.net). The eBURST algorithm v3 was used to assign individual STs to specific CCs (http://eburst.mlst.net).

***spa* typing.** Amplification of the *spa* repeat region was performed using primers *spa* 1113f (5′ AAAGACGATCCTTCGGTGAGC 3′) and *spa* 1514r (5′ CAGCAGTAGTGCCGTTTGCTT 3′) and the conditions described previously (http://www.SeqNet.org). The *spa* types were determined based on the sequencing results using the *spa* plug-in included in the BioNumerics v4.6 software (Applied Math, Sint-Martens-Latem, Belgium).

Genome sequencing. DNA samples were prepared for multiplexed, paired-end sequencing on an Illumina Genome Analyzer IIx (Illumina, Inc., San Diego, CA). For each isolate, 1 to 5 μg DNA in 200 μl was sheared in a 96-well plate with the SonicMAN (part no. SCM1000-3; Matrical BioScience, Spokane, WA) to a size range of 200 to 1,000 bp, with the majority of material at ca. 600 bp, using the following parameters: prechill, 0°C for 75 s; cycles, 20; sonication, 10 s; power, 100%; lid chill, 0°C for 75 s; plate chill, 0°C for 10 s; postchill, 0°C for 75 s. The sheared DNA was purified using the QIAquick PCR Purification kit (catalog no. 28106; Qiagen, Valencia, CA). The enzymatic processing (end repair, phosphorylation, A tailing, and adaptor ligation) of the DNA followed the guidelines described in the Illumina protocol (Preparing Samples for Multiplexed Paired-End Sequencing, catalog no. PE-930-1002, part no.1005361). The enzymes for processing were obtained from New England Biolabs (catalog no. E6000L; New England BioLabs, Ipswich, MA), and the oligonucleotides and adaptors were obtained from Illumina (catalog no. PE-400-1001). After ligation of the adaptors, the DNA was run on a 2% agarose gel for 2 h, after which a gel slice containing 500- to 600-bp

fragments of each DNA sample was isolated and purified using the QIAquick Gel Extraction kit (catalog no. 28706; Qiagen, Valencia, CA). Individual libraries were quantified by quantitative PCR on an ABI 7900HT (part no. 4329001; Life Technologies Corporation, Carlsbad, CA) in triplicate at two concentrations, 1:1,000 and 1:2,000, using the Kapa Library Quantification kit (part no. KK4832 or KK4835; Kapa Biosystems, Woburn, MA). Based on the individual library concentrations, equimolar pools of no more than 12 indexed *S. aureus* libraries were prepared at a concentration of at least 1 nM using 10 mM Tris-HCl (pH 8.0)-0.05% Tween 20 as the diluent. To ensure accurate loading onto the flow cell, the same quantification method was used to quantify the final pools. The pooled paired-end libraries were sequenced on an Illumina Genome Analyzer IIx to a read length of at least 76 bp.

Identification of SNPs. Illumina WGS data sets were aligned against the chromosome of the published ST398 reference genome (strain SO385; GenBank accession no. AM990992) (24) using the short-read alignment component of the Burrows-Wheeler Aligner. Each alignment was analyzed for SNPs using SolSNP (http://sourceforge.net/projects/solsnp/). In order to avoid false calls due to sequencing errors, SNP loci were excluded if they did not meet a minimum coverage of 10× and if the variant was present in less than 90% of the base calls for that position. SNP calls were combined for all of the sequenced genomes such that for the locus to be included in the final SNP matrix, it had to be present in all of the genomes. SNPs falling in the duplicated regions on the reference genome were discarded.

Phylogenetic analysis. Phylogenetic trees were generated using the maximum-parsimony method in PAUP v4.0b10. For maximum-parsimony bootstrapping analysis, the analysis was constrained to build a maximum of 1,000 trees (100 replicates, 10 trees each). The root of the tree was determined through an iterative process as follows. A distance matrix and phylogenetic tree was generated comparing the chromosomes of ST398 (GenBank accession no. AM990992), ST36 (GenBank accession no. BX571856), ST8 (GenBank accession no. CP000255), ST1 (GenBank accession no. BA000033), and ST5 (GenBank accession no. BA000018). Through this process, ST36 was determined to be the most closely related non-CC398 STs. ST36 was used as an outgroup to root the CC398 WGST tree and identify the most ancient CC398 bifurcation point. CC398 descendants nearest to this bifurcation point were used to root subsequent trees.

SCC*mec* typing. The presence of *mecA* and SCC*mec* types and subtypes was assessed in all 89 *S. aureus* CC398 isolates. The structural features unique to each of the type 1 to 5 *ccr* gene complexes; class A, B, and C2 *mec* gene complexes; and four J1 subtypes (a to d) of type IV SCC*mec* were determined by a PCR-based multiplex assay described by Kondo et al. (26). Structural features unique to the class C1 and C1-like *mec* gene complexes (17, 27) and the three subtypes of type V SCC*mec* (17) were determined by aligning the Illumina WGS data sets against reference sequences using CLC Genomics Workbench v4.7.2 (CLC bio, Aarhus, Denmark). The following reference sequences were used: *mec* class C1 (GenBank accession no. AB373032); *mec* class C1-like (GenBank accession no. AB505630); and SCC*mec* subtypes Va (5C2) (GenBank accession no. AB121219), Vb (5C2&5) (GenBank accession no. AB462393), and Vc (5C2&5) (GenBank accession no. AB505629).

SCC*mec* nomenclature was applied as proposed by the International Working Group on the Classification of Staphylococcal Cassette Chromosome Elements (28). For brevity, the type is indicated by roman numerals and the subtype is identified by a lowercase latin letter. The combination of *ccr* and *mec* gene complexes is indicated by an arabic number and a latin letter, respectively, in parentheses. When a composite of two SCC elements carrying distinct *ccr* gene complexes is identified, this is indicated by an ampersand and an arabic numeral designating the *ccr* type.

Detection of genes associated with antimicrobial resistance and host adaptation. All 89 genomes were analyzed for the presence of the tetracycline resistance gene *tet*(M), the cadmium-zinc resistance gene *czrC*, the φSa3 and φSa2 prophages (identified by Sa3int and Sa2int inte-

grase genes), five genes carried by φSa3 prophages (*sea*, *sep*, *sak*, *chp*, and *scn*), two putative avian-niche-specific genes carried by φAvβ (a φSa3-like prophage) (SAAV_2008 and SAAV_2009), and two PVL genes carried by φSa2 prophages (*lukF-PV* and *lukS-PV*). Local BLASTN searches were performed on *de novo* contigs assembled from the Illumina WGS data sets, as well as a reference assembly, using CLC Genomics Workbench v4.7.2. The presence or absence of genes was determined using thresholds of 90% nucleotide identity, 90% coverage of the query sequence length, and a sequence depth of >10×. The query sequences used were *tet*(M) and *czrC* (GenBank accession no. AM990992); Sa3int, *sea*, *sak*, *chp*, and *scn* (GenBank accession no. NC_009641); *sep* (GenBank accession no. BA000018); SAAV_2008 and SAAV_2009 (GenBank accession no. CP001781); and Sa2int, *lukF-PV*, and *lukS-PV* (GenBank accession no. AB006796).

All *de novo* contigs with BLASTN matches to Sa3int were selected, and the Sa3int genes were retrieved for phylogenetic reconstruction using Sa3int (GenBank accession no. NC_009641) and Avβint (GenBank accession no. CP001781) as reference sequences. Gene sequences were aligned using ClustalW v 2.0 (29), and the trees were generated using the maximum-parsimony method in PAUP v4.0b10.

The φSa3 prophages received letter designations to reflect unique combinations of the five prophage-carried genes that modulate human innate immune responses (*sea*, *sep*, *sak*, *chp*, and *scn*) as described elsewhere (30).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00305-11/-/DCSupplemental.

Data set S1, XLSX file, 0.1 MB.
Figure S1, PDF file, 0.3 MB.
Figure S2, PDF file, 0.7 MB.
Figure S3, PDF file, 0.3 MB.
Figure S4, PDF file, 0.7 MB.
Figure S5, PDF file, 0.2 MB.
Figure S6, PDF file, 0.2 MB.
Figure S7a, PDF file, 0.2 MB.
Figure S7b, PDF file, 0.2 MB.

REFERENCES

1. **Devriese LA, Van Damme LR, Fameree L.** 1972. Methicillin (cloxacillin)-resistant *Staphylococcus aureus* strains isolated from bovine mastitis cases. Zentralbl. Veterinaermed. B. **19**:598–605.
2. **Armand-Lefevre L, Ruimy R, Andremont A.** 2005. Clonal comparison of *Staphylococcus aureus* isolates from healthy pig farmers, human controls, and pigs. Emerg. Infect. Dis. **11**:711–714.
3. **Voss A, Loeffen F, Bakker J, Klaassen C, Wulf M.** 2005. Methicillin-resistant *Staphylococcus aureus* in pig farming. Emerg. Infect. Dis. **11**:1965–1966.
4. **Baba K, Ishihara K, Ozawa M, Tamura Y, Asai T.** 2010. Isolation of meticillin-resistant *Staphylococcus aureus* (MRSA) from swine in Japan. Int. J. Antimicrob. Agents. **36**:352–354.
5. **Khanna T, Friendship R, Dewey C, Weese JS.** 2008. Methicillin resistant *Staphylococcus aureus* colonization in pigs and pig farmers. Vet. Microbiol. **128**:298–303.
6. **Lewis HC, et al.** 2008. Pigs as source of methicillin-resistant *Staphylococcus aureus* CC398 infections in humans, Denmark. Emerg. Infect. Dis. **14**:1383–1389.
7. **Smith TC, et al.** 2009. Methicillin-resistant *Staphylococcus aureus* (MRSA) strain ST398 is present in midwestern U.S. swine and swine workers. PLoS One **4**:e4258.
8. **Wulf MW, Verduin CM, van Nes A, Huijsdens X, Voss A.** 2011.
Infection and colonization with methicillin resistant *Staphylococcus aureus* ST398 versus other MRSA in an area with a high density of pig farms. Eur. J. Clin. Microbiol. Infect. Dis. **31**:61–65.
9. **van Cleef BA, et al.** 2011. Livestock-associated methicillin-resistant *Staphylococcus aureus* in humans, Europe. Emerg. Infect. Dis. **17**:502–505.
10. **Hasman H, et al.** 2010. Spa type distribution in *Staphylococcus aureus* originating from pigs, cattle and poultry. Vet. Microbiol. **141**:326–331.
11. **van Belkum A, et al.** 2008. Methicillin-resistant and -susceptible *Staphylococcus aureus* sequence type 398 in pigs and humans. Emerg. Infect. Dis. **14**:479–483.
12. **Gómez-Sanz E, et al.** 2010. Detection, molecular characterization, and clonal diversity of methicillin-resistant *Staphylococcus aureus* CC398 and CC97 in Spanish slaughter pigs of different age groups. Foodborne Pathog. Dis. **7**:1269–1277.
13. **Köck R, et al.** 2009. Prevalence and molecular characteristics of methicillin-resistant *Staphylococcus aureus* (MRSA) among pigs on German farms and import of livestock-related MRSA into hospitals. Eur. J. Clin. Microbiol. Infect. Dis. **28**:1375–1382.
14. **Nübel U, et al.** 2008. Frequent emergence and limited geographic dispersal of methicillin-resistant *Staphylococcus aureus*. Proc. Natl. Acad. Sci. U. S. A. **105**:14130–14135.
15. **Harris SR, et al.** 2010. Evolution of MRSA during hospital transmission and intercontinental spread. Science **327**:469–474.
16. **Pearson T, Okinaka RT, Foster JT, Keim P.** 2009. Phylogenetic understanding of clonal populations in an era of whole genome sequencing. Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis. **9**:1010–1019.
17. **Li S, et al.** 2011. Novel types of staphylococcal cassette chromosome mec elements identified in clonal complex 398 methicillin-resistant *Staphylococcus aureus* strains. Antimicrob. Agents Chemother. **55**:3046–3050.
18. **Graveland H, Wagenaar JA, Bergs K, Heesterbeek H, Heederik D.** 2011. Persistence of livestock associated MRSA CC398 in humans is dependent on intensity of animal contact. PLoS One **6**:e16830.
19. **Foster TJ.** 2005. Immune evasion by staphylococci. Nat. Rev. Microbiol. **3**:948–958.
20. **Rooijakkers SH, et al.** 2005. Immune evasion by a staphylococcal complement inhibitor that acts on C3 convertases. Nat. Immunol. **6**:920–927.
21. **Lowder BV, et al.** 2009. Recent human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*. Proc. Natl. Acad. Sci. U. S. A. **106**:19545–19550.
22. **Cavaco LM, Hasman H, Aarestrup FM.** 2011. Zinc resistance of *Staphylococcus aureus* of animal origin is strongly associated with methicillin resistance. Vet. Microbiol. **150**:344–348.
23. **Robinson DA, Enright MC.** 2004. Evolution of *Staphylococcus aureus* by large chromosomal replacements. J. Bacteriol. **186**:1060–1064.
24. **Schijffelen MJ, Boel CH, van Strijp JA, Fluit AC.** 2010. Whole genome analysis of a livestock-associated methicillin-resistant *Staphylococcus aureus* ST398 isolate from a case of human endocarditis. BMC Genomics **11**:376.
25. **Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG.** 2000. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. J. Clin. Microbiol. **38**:1008–1015.
26. **Kondo Y, et al.** 2007. Combination of multiplex PCRs for staphylococcal cassette chromosome mec type assignment: rapid identification system for mec, ccr, and major differences in junkyard regions. Antimicrob. Agents Chemother. **51**:264–274.
27. **Berglund C, et al.** 2008. Novel type of staphylococcal cassette chromosome mec in a methicillin-resistant *Staphylococcus aureus* strain isolated in Sweden. Antimicrob. Agents Chemother. **52**:3512–3516.
28. **Iwg SCC.** 2009. Classification of staphylococcal cassette chromosome mec (SCC*mec*): guidelines for reporting novel SCC*mec* elements. Antimicrob. Agents Chemother. **53**:4961–4967.
29. **Larkin MA, et al.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23**:2947–2948.
30. **van Wamel WJ, Rooijakkers SH, Ruyken M, van Kessel KP, van Strijp JA.** 2006. The innate immune modulators staphylococcal complement inhibitor and chemotaxis inhibitory protein of *Staphylococcus aureus* are located on beta-hemolysin-converting bacteriophages. J. Bacteriol. **188**:1310–1315.