Identification of a Highly Transmissible Animal-Independent *Staphylococcus aureus* ST398 Clone with Distinct Genomic and Cell Adhesion Properties

Anne-Catrin Uhlemann,[a] Stephen F. Porcella,[b] Sheetal Trivedi,[a] Sean B. Sullivan,[a] Cory Hafer,[a] Adam D. Kennedy,[c] Kent D. Barbian,[b] Alex J. McCarthy,[d] Craig Street,[e] David L. Hirschberg,[e] W. Ian Lipkin,[e] Jodi A. Lindsay,[d] Frank R. DeLeo,[c] and Franklin D. Lowy[a]

Department of Medicine, Division of Infectious Diseases, College of Physicians and Surgeons, Columbia University, New York, New York, USA[a]; Genomics Unit, Research Technologies Section,[b] and Laboratory of Human Bacterial Pathogenesis,[c] Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, Montana, USA; Centre for Infection, Division of Clinical Sciences, St. George's, University of London, London, United Kingdom[d]; and Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, New York, USA[e]

ABSTRACT **A methicillin-resistant *Staphylococcus aureus* (MRSA) clone known as ST398 has emerged as a major cause of acute infections in individuals who have close contact with livestock. More recently, the emergence of an animal-independent ST398 methicillin-sensitive *S. aureus* (MSSA) clone has been documented in several countries. However, the limited surveillance of MSSA has precluded an accurate assessment of the global spread of ST398 and its clinical relevance. Here we provide evidence that ST398 is a frequent source of MSSA infections in northern Manhattan and is readily transmitted between individuals in households. This contrasts with the limited transmissibility of livestock-associated ST398 (LA-ST398) MRSA strains between humans. Our whole-genome sequence analysis revealed that the chromosome of the human-associated ST398 MSSA clone is smaller than that of the LA-ST398 MRSA reference strain S0385, due mainly to fewer mobile genetic elements (MGEs). In contrast, human ST398 MSSA isolates harbored the prophage φ3 and the human-specific immune evasion cluster (IEC) genes *chp* and *scn*. While most of the core genome was conserved between the human ST398 MSSA clone and S0385, these strains differed substantially in their repertoire and composition of intact adhesion genes. These genetic changes were associated with significantly enhanced adhesion of human ST398 MSSA isolates to human skin keratinocytes and keratin. We propose that the human ST398 MSSA clone can spread independent of animal contact using an optimized repertoire of MGEs and adhesion molecules adapted to transmission among humans.**

IMPORTANCE ***Staphylococcus aureus* strains have generally been considered to be species specific. However, cross-species transfers of *S. aureus* clones, such as ST398 methicillin-resistant *S. aureus* (MRSA), from swine to humans have been reported. Recently, we observed the emergence of ST398 methicillin-susceptible *S. aureus* (MSSA) as a colonizing strain of humans in northern Manhattan. Here we report that ST398 is a frequent cause of MSSA infections in this urban setting. The ST398 MSSA clone was readily transmitted within households, independent of animal contact. We discovered that human ST398 MSSA genomes were smaller than that of the LA-ST398 strain S0385 due to fewer mobile genetic elements. Human and LA-ST398 strains also differed in their composition of adhesion genes and their ability to bind to human skin keratinocytes, providing a potential mechanism of *S. aureus* host adaptation. Our findings illustrate the importance of implementing molecular surveillance of MSSA given the evidence for the rapid and clinically undetected spread of ST398 MSSA.**

Received 30 January 2012 **Accepted** 3 February 2012 **Published** 28 February 2012

Citation Uhlemann A, et al. 2012. Identification of a highly transmissible animal-independent *Staphylococcus aureus* ST398 clone with distinct genomic and cell adhesion properties. mBio 3(2):e00027-12. doi:10.1128/mBio.00027-12.

Editor Donald Low, Mt. Sinai Hospital

Address correspondence to Anne-Catrin Uhlemann, au2110@columbia.edu.

Staphylococcus aureus is a highly adaptable commensal organism and an important pathogen of humans and animals (1, 2). Molecular epidemiology studies suggest that genetic subsets of *S. aureus* lineages are particularly well suited to causing infections in certain animal species (3, 4). Recently, swine have been identified as a major reservoir for methicillin-resistant *S. aureus* (MRSA) (5). Most of these MRSA strains belong to the multilocus sequence type (MLST) 398 (ST398), are nontypeable by SmaI restriction in pulsed-field gel electrophoresis (6), and can be identified by an ST398-specific restriction-modification (RM) test (7).

Importantly, ST398 MRSA strains are also responsible for a significant number of acute infections, primarily in individuals who have direct contact with animals (8). ST398 was first recognized in France (8) and has been noted to cause a substantial number of MRSA infections in The Netherlands and other parts of Europe (9). This strain has now also been detected in China (10) and North America (11). The spectrum of disease most commonly includes superficial skin and soft tissue infections; however, invasive infections, such as pneumonia (12), bacteremia, and endocarditis, have also been reported (13).

TABLE 1 Presence of ST398 in community households in northern Manhattan[a]

Characteristic	Value for households with:		*P* value
	ST398 present	ST398 absent[b]	
No. of households with epidemiological risk factor/no. in group (%)			
Presence of pet in house	6/19 (31.6)	86/303 (28.4)	0.764
Shopping at live poultry market	8/19 (42.1)	81/303 (26.7)	0.146
International travel index, past 6 mo	5/19 (26.3)	62/303 (20.5)	0.561
Transmission dyads			
No. of households member dyads	42	915	
No. of concordant dyads/no. in group (%)	8/42 (19.1%)	90/915 (9.8%)	0.065

[a] Dyads, household member pairs colonized with the same clone.
[b] For households in which ST398 is absent, dyad data refer to all other strains.

To date, one of the defining epidemiological features of livestock-associated ST398 (LA-ST398) MRSA is the correlation between human colonization and the nature and duration of contact with animals (14, 15). In local communities, the spread of LA-ST398 MRSA beyond the immediate animal-exposed family has rarely been observed and persistent nasal colonization is infrequently detected in individuals without direct animal exposure (15). In addition, although nosocomial transmission and ventilator-associated infections with LA-ST398 MRSA have been described (16), this clone appears to be less transmissible than hospital-associated MRSA (HA-MRSA) (17).

Recently, we observed the emergence of ST398 methicillin-sensitive *S. aureus* (MSSA) in community households in northern Manhattan (18). Colonized individuals did not report animal contacts or recent travel, but this largely Dominican population has close links to the Dominican Republic, where this strain was also detected, raising speculation that it may have been imported. However, an increasing number of case reports have described ST398 MSSA infections in China and in parts of Europe, such as France and The Netherlands, including cases of necrotizing pneumonia and invasive bloodstream infections in young healthy individuals (13, 19). The concern that these strains perhaps represent a more virulent ST398 subtype is further supported by a Dutch surveillance study. There, only 0.2% of healthy individuals were colonized by ST398 MSSA, whereas ST398 was recovered at a higher frequency from bloodstream infections (2.1%) (13). Notably, most of those human ST398 MSSA infections were acquired in the absence of documented exposure to livestock. These observations prompted speculations that variants of the ST398 lineage may persist in humans without direct animal contact.

Genotyping of human ST398 MSSA isolates by *S. aureus* protein A (*spa*) typing indicates that most of these infections are caused by *spa* type t571 strains, whereas the predominant *spa* types among pig-associated isolates are t011, t034, and t108. Recently, Schijffelen et al. reported the genome sequence of a LA-ST398 MRSA strain (labeled S0385) isolated from a case of human endocarditis (20). This strain contained considerable differences in its accessory genome content relative to other *S. aureus* genomes, including the presence of mobile genetic elements (MGEs) that confer antibiotic resistance, multiple integrative conjugative elements (ICEs), and a unique *S. aureus* pathogenicity island (SaPI). Several virulence factors, such as enterotoxins and phage-encoded toxins, were lacking in this isolate. Analysis of additional pig- or human-derived ST398 MRSA isolates yielded an inconsistent presence of these MGEs (20).

Here we have used a combination of epidemiological and molecular approaches to query the basis for success of the northern Manhattan ST398 MSSA clone (ST398-NM) as a colonizer and pathogen in humans who do not have contact with livestock.

RESULTS

ST398 MSSA isolates are frequently recovered from patients with *S. aureus* infections and from nasal swabs of healthy individuals in northern Manhattan. To examine the incidence of ST398 infections in the northern Manhattan community, we *spa* typed 160 consecutive clinical noninvasive MSSA isolates, 161 randomly collected clinical outpatient MRSA isolates, and 160 bloodstream MSSA infections. Eight of the 160 (5%) noninvasive MSSA isolates were identified as ST398 (*spa* type t571) and were among the most common MSSAs, along with the USA300 t008 ($n = 9$ [5.6%]), ST5 t002 ($n = 13$ [8.1%]), and ST30 t665 ($n = 19$ [11.9%]) strains. These MSSA ST398 isolates were derived primarily from skin and soft tissue infections ($n = 4$). None of the MRSA isolates were found to be ST398. ST398 t571 was also detected among the MSSA bloodstream isolates (4/160 [2.5%]). All ST398 MSSA isolates were resistant to erythromycin and clindamycin but, in contrast to reports of LA-ST398 (21), remained sensitive to tetracycline.

To further investigate potential community reservoirs of ST398 MSSA in northern Manhattan (18), we analyzed the prevalence of ST398 colonization among household members or on environmental surfaces as part of an ongoing study of *S. aureus* transmission in 332 community households. We detected *S. aureus* colonization in 232 households. Of these, 19 (8.2%) households harbored ST398, either as a human colonizer or on environmental surfaces (14 or 15 households, respectively). The presence of ST398 in a household was not associated with having a pet, shopping at a live poultry market, or recent travel (Table 1). To estimate the spread of *S. aureus* clones, we identified household member pairs (dyads) colonized with the same clone (Table 1). In households with ST398 colonization, we detected 8 ST398-positive member dyads out of 42 possible dyads (19.1%). By comparison, there were fewer concordant dyads among all other *S. aureus* clones (90/915 [9.8%]), including USA300 (15/171 [8.8%]). Taken together, these results indicate that ST398 MSSA is efficiently spreading independent of animal contact among people living in northern Manhattan.

ST398-NM genome organization and MGEs. To elucidate the molecular basis of the epidemiological success of ST398-NM, we sequenced the genomes of two representative isolates (ST398NM01 and ST398NM02) and compared these genomes to

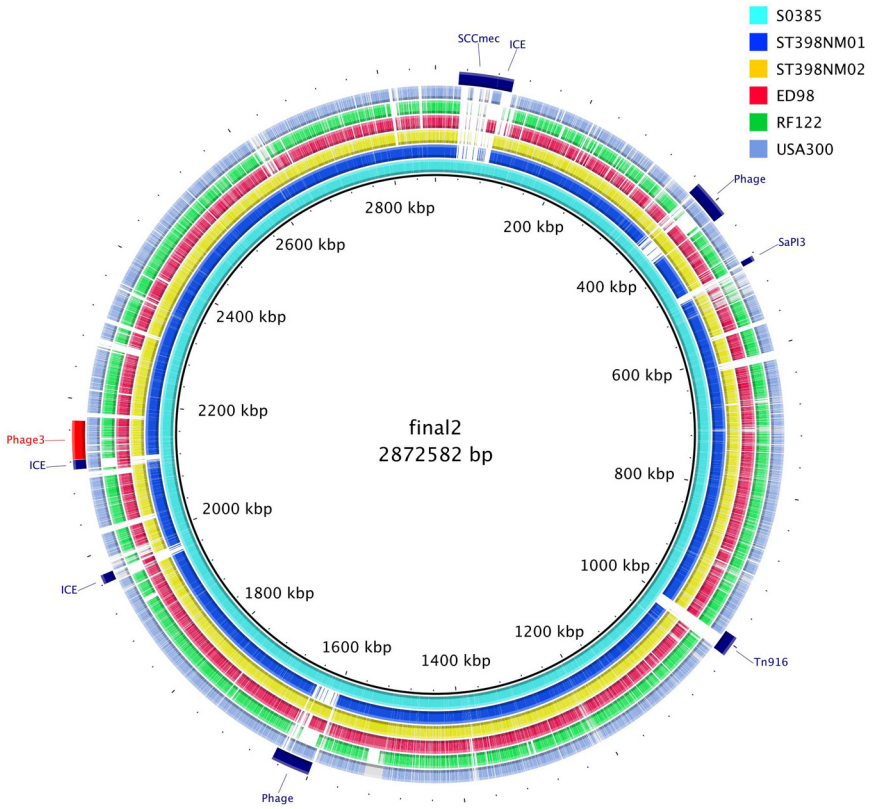

FIG 1 Comparison of *S. aureus* genomes. From the inside: S0385 (aqua, inner ring), ST398NM01 (blue), ST398NM02 (yellow), ED98 (red), RF122 (green), and USA300 (light-blue) genomes are depicted. In the outer circle, features of S0385 and ST398NM01 are displayed in navy and red, respectively.

the only fully annotated ST398 genome (labeled S0385), a LA-ST398 MRSA isolate (20).

ST398NM01 was the first identified colonizing isolate in 2004 (18). Its genome consists of a circular chromosome of 2,714,439 bp and is 170,928 bp smaller than that of S0385 (see Table S1 in the supplemental material). This difference in size is mainly accounted for by considerably fewer MGEs being present in ST398NM01, since this isolate lacks SCCmec, SaPI5, the three ICEs, one transposon, and both prophages (Fig. 1; see also Table S1 in the supplemental material). In contrast, ST398NM01 carries a variant of the prophage φ3, containing the immune evasion complex (IEC) genes encoding the chemotaxis inhibitory protein (*chp*) and staphylococcal complement inhibitor (*scn*) but not that encoding staphylokinase (*sak*). This prophage is integrated into the β-hemolysin locus and has been found in most isolates infecting humans but not animals (22, 23). In addition, a small plasmid (~3.9 kb) was inserted into the ST398NM01 genome. This plasmid contains genes encoding the cadmium efflux system accessory protein (*cadX*), the cadmium resistance protein (*cadD*), and the rRNA adenine *N*-6-methyltransferase (*ermT*), consistent with the observed erythromycin resistance in ST398 MSSA.

Following recent reports of fulminant staphylococcal infections with Panton-Valentine leukocidin (PVL)-containing ST398 (t571) strains in France, we also performed a draft genome survey of a PVL-positive abscess isolate (ST398NM02) collected in the Dominican Republic to determine the presence of yet-unidentified virulence factors in the ST398 lineage (24). This genome sequence is currently at 40 contigs and comprises 2,725,515 bp (see Table S1 in the supplemental material). Most of the gaps are located in rRNA regions or within repetitive sequences. ST398NM02 also contains prophage φ3, which encodes the IEC genes *chp* and *scn*, and prophage φ2, which encodes the PVL LukS and LukF subunits (*lukS-PV* and *lukF-PV*). This isolate does not carry an integrated plasmid and also lacks any additional novel MGEs.

Both human ST398 MSSA genomes lack the Tn*916*-like transposon encoding tetracycline resistance in strain S0385, but they contain an allelic variant of the Tn*552* transposon, encoding an inducible β-lactamase and its regulatory components, as well as the Tn*7*-like transposon described for S0385 (20).

Human ST398-NM and LA-ST398 strain S0385 carry different MGEs. To further determine the prevalence of S0385 MGEs among geographically diverse ST398 isolates (20), we screened for the presence of MGEs in our entire collection of ST398 MSSA clinical isolates from northern Manhattan, the Dominican Republic, Martinique, and France ($n = 48$) and in LA-ST398 MRSA isolates from Canada ($n = 5$) by PCR. None of the human ST398 MSSA isolates carried SaPI5-located variants of *vwb* or *scn* or genes typically present in prophage φ6 or ICE2. These genes were present at low frequencies in Canadian LA-ST398 isolates (Table 2) (20), whereas the tetracycline resistance gene *tetM* was present in all pig ST398 isolates but in none of the human isolates (Table 2).

To further investigate if the gene content in ST398 MSSA isolates differed by infection status, time of collection, or geographic

TABLE 2 PCR screening for S0385 mobile genetic elements in geographically diverse ST398 isolates

Strain collection (*n*)[a]	No. (%) of isolates with:								
		Mobile genetic element[b]							
	spa type t571	*vwb*	*scn*	*pvl*	Phage φ2	Phage φ6	ICE1	ICE2	*tetM*
Animal									
Pig (NL) (7)	0 (0)	2 (29)	2 (29)	NR[c]	5 (71)	2 (29)	1 (14.3)	1 (14.3)	7 (100)
Pig (CAN) (5)	1 (20)	2 (40)	4 (80)	0 (0)	0 (0)	5 (100)	3 (60)	0 (0)	5 (100)
Human									
NYC (22)	17 (77)	0 (0)	0 (0)	2 (9)	2 (9)	0 (0)	8 (36)	0 (0)	0 (0)
France (11)	6 (55)	0 (0)	0 (0)	1 (9)	2 (18)	0 (0)	2 (18)	0 (0)	0 (0)
DR (7)	7 (100)	0 (0)	0 (0)	2 (29)	2 (29)	0 (0)	4 (57)	0 (0)	0 (0)
MQ (8)	5 (63)	0 (0)	0 (0)	0 (0)	1 (13)	0 (0)	0 (0)	0 (0)	0 (0)

[a] *n*, samples tested. NL, The Netherlands; CAN, Canada; NYC, New York City; DR, Dominican Republic; MQ, Martinique.
[b] *vwb* and *scn* are SaPI variants.
[c] NR, not reported.

origin, we used a 62-strain custom *S. aureus* microarray (SAM-62) (25) to perform comparative genome hybridization (CGH) analysis of eight selected human ST398 clinical or colonizing isolates from northern Manhattan, the Dominican Republic, and Martinique (18, 24, 26). All isolates carried prophage φ3, which contains *chp* and *scn* only, and nearly half of the strains contained an additional prophage (φ1, φ2, or φ5) (Table 3; see also Fig. S1 in the supplemental material). All but one of the isolates harbored *cadX*, consistent with the presence of the integrated plasmid identified in ST398NM01. Overall, there was remarkably little variation in MGEs in this diverse collection of human-colonizing and infectious ST398 MSSA isolates, which suggests recent clonal expansion and dissemination of a human ST398 lineage.

Common features of ST398 MSSA and MRSA isolates. Despite the considerable variation in MGEs and the core genome of human ST398 isolates and LA-ST398 strain S0385, all three genomes share a number of unique features. They contain the previously described allelic variant of a unique type I restriction modification (RM) *hsdS* gene but lack a second copy of a type I RM system. In contrast, they harbor a putative type II RM system, encoding a type II endonuclease (SsoII) and a novel cytosine-specific DNA methylase (EcoRII). This methylase is unique among *S. aureus* genomes and likely accounts for the characteristic SmaI resistance of the ST398 clonal lineage. Phylogenetically, the closest relative of this methylase is found in *S. pseudintermedius* (see Fig. S2 in the supplemental material). This zoonotic pathogen, sometimes associated with disease in humans, is known mainly for being a colonizer and pathogen in dogs.

Comparison of ST398 core genomes. Compared with the core genome (i.e., excluding MGEs) of S0385, there were 381 single nucleotide polymorphisms (SNPs) in 272 genes in ST398NM01 (247 nonsynonymous [dN] and 134 synonymous [dS] SNPs; dN/dS ratio = 1.8) and 356 SNPs in ST398NM02 (234 nonsynonymous SNPs, 122 synonymous SNPs; dN/dS ratio = 1.9). Strains ST398NM01 and ST398NM02 differed from each other by 121 SNPs (81 nonsynonymous and 40 synonymous SNPs), consistent with a phylogenetically closer relationship of the two human ST398 isolates compared with the LA-ST398 strain.

To determine if SNPs clustered with significant excess by cellular processes, we carried out BLAST searches to analyze SNP distribution among clusters of orthologous groups of proteins (COGs) (27). The COG database classifies proteins by their evolutionary relationship and allows further assignment of proteins into functional categories. In ST398NM01, SNPs could be detected in a wide range of different COGs (27) but were observed most frequently in genes encoding proteins involved in amino acid metabolism and transport ($n = 30$) or general functional prediction ($n = 31$) or genes not clustering into COGs ($n = 67$). A high ratio of nonsynonymous-to-synonymous SNPs (dN/dS > 1) among closely related bacteria is thought to indicate recent diversification (28), since purifying selection has not had sufficient time to remove deleterious substitutions (29). The highest dN/dS ratios were observed for genes involved in coenzyme metabolism (dN/dS ratio = 7), in conserved genes of unknown function (dN/dS ratio = 5), and in genes encoding amino acid metabolism, translation, transcription, lipid metabolism, or cell wall synthesis (see Table S2 in the supplemental material). The relatively high dN/dS ratios for genes within selected COGs and metabolic pathways could be evidence of recent niche or host adaptation of ST398.

There was noted variation in genes encoding cell surface-bound proteins in the human ST398 genomes relative to that of

TABLE 3 Microarray analysis of geographically distinct ST398 MSSA isolates

Isolate	Origin[a]	Clinical context	*spa* type	Bacteriophage	Plasmid *rep* gene(s)	Virulence genes	Resistance genes
DR10	DR, 2007	Abscess	t571	φ2, φ3	rep_{25}	*PVL*, *chp*, *scn*	*blaZ*, *smr*
DR122	DR, 2008	Colonizer	t571	φ3, φ5	rep_{25}, rep_{27}, rep_{30}	*chp*, *scn*	*blaZ*, *cadDX*
50612	NYC 2010	Colonizer	t5635	φ3	rep_{25}, rep_{27}	*chp*, *scn*	*blaZ*, *cadDX*, *smr*
51246	NYC 2010	Colonizer	t571	φ3	rep_{27}	*chp*, *scn*	*blaZ*, *cadDX*
S42	MQ, 2008	Wound	t571	φ3	rep_{27}	*chp*, *scn*	*cadDX*
71633	NYC 2004	Colonizer	t571	φ3	rep_{27}	*chp*, *scn*	*blaZ*, *cadDX*
MS298	NYC 2010	Wound	t1451	φ1, φ3	rep_{25}, rep_{27}	*chp*, *scn*	*blaZ*, *cadDX*
MS636	NYC 2011	Wound	t571	φ3	rep_{13}, rep_{27}	*chp*, *scn*	*blaZ*, *cadDX*, *smr*

[a] Place of origin, yr. DR, Dominican Republic; MQ, Martinique; NYC, New York City/northern Manhattan.

TABLE 4 Variation of ST398 surface proteins[a]

Gene	Description[b]		
	Human ST398		
	ST398NM01	ST398NM02	LA-ST398 (S0385)
clfA	WT	WT	Truncated
clfB	WT	Δ456 bp	Truncated
sdrC	WT	WT	Δ174 bp
sdrD	Δ54 bp	WT	3 SNPs
sdrE	WT	Δ138 bp	Absent
sdrH	Δ39 bp	WT	WT
fnbA	WT	Δ42 bp	WT
fnbB	WT	WT	Truncated
coa	WT	WT	Δ81 bp
cna	ΔB domain	ΔB domain	WT
isdB	WT	WT	SNP, Δ9 bp
ebpS	WT	WT	WT
ecb	WT	WT	WT
Gene encoding hypothetical protein	WT	WT	Δ174 bp

[a] A table with gene locus tags is in the supplemental material.
[b] WT, wild type.

strain S0385 (Table 4), such as in the gene encoding the iron-responsive surface determinant IsdB and deletions or insertions within repetitive sequence regions of the genes encoding serine aspartate repeat proteins (*sdrC* and *sdrH*), the large erythrocyte membrane binding protein (*emb*), and the collagen adhesin (*cna*; deletion of B domain [Table 4]).

The LA-ST398 strain S0385 lacks *sdrE*, which is known to be absent from a number of *S. aureus* genomes (30). In addition, genes encoding clumping factors A and B (*clfA* and *clfB*), as well as that encoding fibronectin-binding protein B (*fnbB*), harbored premature stop codons, consistent with evidence of gene decay in strain S0385 (Table 4). There were no truncations of *clfA*, *clfB*, or *fnbB* in the five Canadian LA-ST398 isolates tested, but they harbored domain deletions or insertions in variable repeat regions that were distinct from the human ST398 isolates.

In contrast, genes encoding a number of surface-bound proteins were conserved between human ST398 genomes and strain S0385, such as those encoding fibronectin-binding protein A (*fnbA*), elastin-binding protein of *S. aureus* (*ebpS*), and the extracellular complement-binding protein (*ecb*) (Table 4). Among known secreted surface proteins, only the coagulase gene (*coa*) was found to have an 81-bp deletion in S0385, whereas all other proteins in this class were preserved. The accumulation of truncations and other nucleotide differences in genes encoding select surface adhesion proteins in strain S0385 relative to sequence of the human ST398 isolates suggests differential adaptation to host extracellular matrix proteins.

Human ST398 MSSA isolates have increased binding to keratinocytes and purified keratin compared to LA-ST398 MRSA isolates. The ability of *S. aureus* to interact with keratinocytes and nasal squamous epithelial cells is important for human colonization and infection. We hypothesized that the differences in transmissibility (or perhaps species tropism) of human ST398 MSSA and LA-ST398 MRSA may be a reflection of the allelic variation in surface adhesion proteins (e.g., MSCRAMMs) between these strain types and their ability to interact with host receptors. Using *in vitro* binding assays, we observed significantly increased adherence of human ST398 MSSA isolates ($n = 5$) to normal human epidermal keratinocytes (NHEK) compared to LA-ST398 MRSA isolates ($n = 5$) ($P = 0.004$) (Fig. 2). Human ST398 isolates were also more adherent to type I keratin meta-keratin 4 and to the ClfB-binding domain of recombinant cytokeratin 10 (31) (Fig. 2). In contrast, human and LA-ST398 isolates did not differ in their binding to primary porcine keratinocytes (Fig. 2). These *in vitro* studies illustrate the potential importance of the adhesion genes in the colonization and transmission of the ST398 MSSA clone among humans.

DISCUSSION

In contrast to LA-ST398 MRSA, where the spread of the clone is limited to farm personnel, the ST398 MSSA clone found in northern Manhattan is readily transmitted among humans independent of animal contact. This ST398 MSSA clone is now encountered among humans in geographically dispersed areas (10, 13, 18, 19).

Here we present epidemiological, molecular, and comparative genomic evidence that MSSA isolates of the ST398 lineage are easily transmissible among humans and have a genome that is well adapted to the human host. Specifically, we found that ST398-NM causes a wide range of infections, including invasive bloodstream infections, and is a frequent colonizer within households. These attributes are remarkably different from those of LA-ST398 MRSA, for which transmission is limited to farmers, veterinarians, and their immediate household members (32). The key genotypic and phenotypic features that distinguish both ST398-NM strains from the LA-ST398 strain S0385 include the following: (i) a different repertoire of MGEs, (ii) accessory genome variation, including a complete repertoire of surface adhesins in the human ST398 strains as opposed to pseudogenes in strain S0385, (iii) an increased ability of human ST398 to adhere to human keratinocytes and immobilized keratin, and (iv) clustering of SNPs in selected functional COGs in the core genomes.

These differences at the genome level suggest several mechanisms of ST398 host-specific adaptation. Both human ST398 MSSA genomes lacked novel toxins or an *S. aureus* SaPI, but their repertoire of MGEs was most consistent with those of other known human *S. aureus* genomes, such as prophage φ3 being integrated into the β-hemolysin locus or the presence of the PVL-

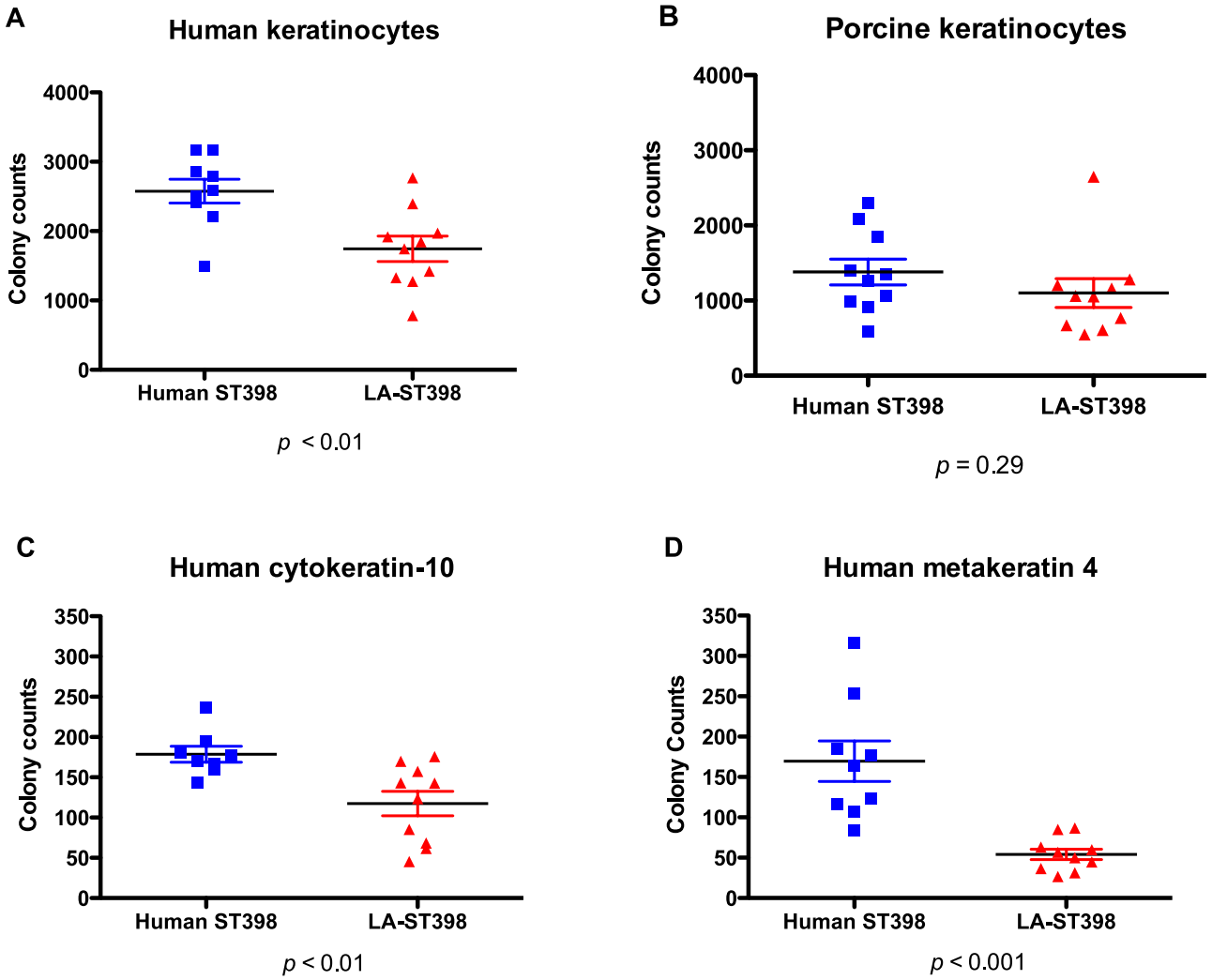

FIG 2 Binding of human ST398 MSSA isolates (blue squares) and LA-ST398 MRSA isolates (red triangles) to NHEK cells ($P = 0.0043$) (A), porcine keratinocytes ($P = 0.29$) (B), the ClfB-binding domain of cytokeratin 10 ($P = 0.0057$) (C), or meta-keratin 4 ($P = 0.0003$) (D). Data are colony counts ± SEM from two independent experiments performed in triplicate.

encoding prophage φ2. Furthermore, there was very limited variation in MGEs among a geographically diverse collection of colonizing and infectious ST398 MSSA isolates. This may be consistent with the idea that there is relatively little pressure to rapidly adapt to new host niches or might indicate recent emergence and rapid dissemination of human ST398.

In contrast, the LA-ST398 strain S0385 contains several unique MGEs, including SaPI5, encoding allelic variants of the von Willebrand factor-binding protein (*vWbp*) and staphylococcal complement inhibitor (*scn*). Related ruminant and equine vWbps have recently been shown to contain a unique N-terminal region specific for the activation of ruminant and equine prothrombin, thereby contributing to host specificity (33). However, SaPI5 is absent from many pig ST398 isolates, therefore providing support to the idea that these genes are not essential for pig specificity (29). It should be noted that S0385 was isolated from a case of human endocarditis in an immunocompromised host, and therefore the acquisition of some of the MGEs in this isolate may have occurred during human infection.

Differences between LA-ST398 and human ST398-NM were not limited to MGEs. The observed gene decay in adhesion proteins in strain S0385 suggests that these genetic factors are likely no longer contributing to virulence in swine and further imply that the most recent common ancestor of the ST398 lineage was human associated. This observation is consistent with studies of *S. aureus* isolates causing disease in poultry, where a number of genes involved in pathogenesis of human *S. aureus* infections, such as *S. aureus* protein A (Spa), were rendered nonfunctional (34). Accumulation of pseudogenes within predicted surface-expressed structural and adhesion proteins has also been noted in the bovine mastitis strains ET3-1 and ET3-2 (lineage CC151), including in *spa* and *clfA* (35), but not in the clinically less successful strain ET3-3. In the current study, we found that human ST398 MSSA isolates bound more avidly to human keratinocytes and keratin derivatives than pig-derived ST398 isolates. We suggest that this differential binding ability may translate directly into decreased transmission of LA-ST398 among humans lacking close animal contact. In our study, binding of LA-ST398 isolates to

human keratinocytes was diminished, but it was not abolished. We propose that this ability to adhere to human cells may be sufficient to initiate transmission from livestock to humans, but it would not be sufficient to maintain spread between humans. The LA-ST398 strain S0385 appears to lack many of the established *S. aureus* virulence factors, an attribute that probably limited human infection to an immunocompromised host (20). This is likely a reflection of the opportunistic nature of *S. aureus* and of the importance of the host response in determining susceptibility to infections. Taken together, our findings support host-specific adaptation of human- and animal-tropic *S. aureus* strains by accumulation of MGEs and mutation of surface-expressed *S. aureus* proteins.

We also note that human ST398 MSSA isolates were able to adhere to primary porcine keratinocytes *in vitro*, and this finding may indicate that human *S. aureus* strains in general are able to initiate colonization in exposed livestock. However, the factors that determine the unique success of the ST398 MRSA lineage as colonizers with little disease in livestock remain incompletely understood. It has been speculated that the presence of two tetracycline resistance determinants (*tetM* and *tetK*) is responsible for the selection of LA-ST398 MRSA and a reflection of the selective pressure exerted by tetracycline antibiotic use in animal feeds (36). In contrast, all human isolates tested here lacked tetracycline resistance, but virtually all carried *ermT*, consistent with macrolide resistance. The importance of these resistance determinants for ST398 strain virulence and fitness remains to be determined.

The presence of ST398 MSSA (mainly *spa* type t571) as a source of infections in geographically dispersed regions, including China, many European countries, North America, the Caribbean, and Colombia, supports the notion that MSSA strains constitute potentially pandemic pathogens. Our global knowledge about the molecular epidemiology of MSSA is limited, since much of the surveillance work on *S. aureus* is centered on MRSA strains. These findings highlight the importance of expanding our knowledge of the molecular epidemiology of MSSA strains globally.

Taken together, our integrated approach of epidemiology and whole-genome analysis of the ST398 lineage has revealed that ST398-NM MSSA constitutes a readily transmissible and clinically important clone that differs significantly at the genome level from its livestock-associated counterpart.

MATERIALS AND METHODS

Ethics statement. We obtained written informed consent from each individual before conducting an interview or obtaining samples. We received parental consent for participating children <18 years old, and pediatric assent was obtained from those capable of providing it. Index participants were compensated $10 for their time. The Institutional Review Board of Columbia University Medical Center (CUMC), New York, NY, approved this study.

Assessment of molecular epidemiology of ST398 in northern Manhattan. We collected 160 consecutive MSSA isolates from outpatients residing in the northern Manhattan neighborhood as defined by zip codes and 160 bloodstream MSSA isolates processed and archived by the Clinical Microbiology Laboratory of the CUMC. We also analyzed 161 MRSA clinical isolates from outpatients (26) who are part of an ongoing case-control cohort study of the transmission of community-associated MRSA (CA-MRSA) in northern Manhattan. One hundred sixty-one patients with MRSA infections presenting from the community along with age-matched controls attending the hospital's dental clinic and their household members were enrolled. Interview data regarding risk factors for *S. aureus* were obtained from each index participant. To assess the degree of strain similarity within *S. aureus* households, concordant *spa* type dyads between household members were analyzed. Anterior nares cultures were collected with sterile premoistened swabs (Becton Dickinson) from index participants and consenting household members.

All statistical analyses were performed using the SPSS 18 software program. Chi-square tests were used for comparison of dichotomous variables, and Fisher's exact test was used with an expected cell count of <5. A P value of <0.05 was considered statistically significant.

Genome sequencing, annotation, and comparative analysis. The first ST398 MSSA strain isolated in 2004, designated 71193 or ST398NM01, was selected for complete genome sequencing (18). DNA sequencing was performed using fragment and paired-end libraries on a Roche genome sequencer Titanium instrument (454 Life Sciences [a Roche company], Branford, CT) to 40× coverage or higher according to the manufacturer's recommendations. Reads were assembled using the gsAssembler software program into 35 contigs, and the order of contigs was determined by alignment to the S0385 genome. Strain DR 10 or ST398NM02 was obtained from the gluteal abscess of a 3-year-old girl in Santo Domingo (Dominican Republic) (24) and was chosen for genome sequencing. DNA sequencing was performed using fragment libraries run on a Roche genome sequencer (454 Life Sciences [a Roche company], Branford, CT) to greater than 40× coverage. Reads were assembled using gsAssembler into 44 contigs. The order of these contigs was determined by alignment to the S0385 genome.

All gaps between contigs in ST398NM01 and selected gaps in ST398NM02 were closed by primer design, PCR fragment generation, Sanger sequencing of the PCR products, and primer walking when needed. Open reading frame (ORF) calling was performed using public and proprietary algorithms, with a minimum length cutoff of 40 amino acids, as previously described (37, 38). The genome sequence and annotation of ST398NM01 and ST398NM02 are deposited in DDBJ/EMBL/GenBank under the NCBI accession numbers CP003045 and AIDT00000000. ORFs displaying evidence of frameshifts or mutations leading to premature stop codons were identified by proprietary algorithms and were manually verified. Genome comparisons were carried out using ClustalW alignments. LA-ST398 MRSA strain S0385 (NCBI AM990992), bovine strain RF122/ET-3 (NCBI AJ938182), avian strain ED98 (NCBI CP001781), and the human epidemic USA300 clone FPR3757 (NCBI CP000255) were used as reference genomes. The BLAST Ring Image Generator (BRIG) was used for visualizing genome alignments (39).

Screening for virulence and resistance genes. The detection of Panton-Valentine leukocidin (PVL) (40) and screening for selected virulence and resistance genes carried by mobile genetic elements in strain S0385 was performed by PCR as described (20). Eleven MSSA isolates from human ST398 infections in France were kindly provided by Frederic Laurent, and five colonizing isolates from pigs in Ontario, Canada, were kindly provided by Scott Weese.

Microarray-based comparative genome analysis. Microarray experiments were performed with 8 ST398 MSSA isolates using a 62-strain *S. aureus* microarray (SAM-62), as previously described (29). This array harbors 29,739 60-mer oligonucleotide probes that represent 6,520 genes and an additional 579 gene variants (25). Importantly, 6 genomes of animal isolates are represented on SAM-62 (2 cow, 1 sheep, 1 chicken, and 2 pig), including 2 clonal complex 398 (CC398) isolates. We performed hierarchical clustering analysis using the Euclidean distance metric based on 18,024 60-mer oligonucleotides representing genes found in the core and CV genomes.

Cell cultures and binding assays. Normal human epidermal keratinocytes (NHEK) derived from neonatal foreskin keratinocytes were kindly provided by the Department of Dermatology at Columbia University. Cells were maintained serum free in the keratinocyte basal medium KBM-Gold (Lonza, Walkersville, MD) supplemented with KGM-Single Quot (Lonza) and grown in 12-well plates to >90% confluence. Primary por-

cine keratinocytes were obtained by trypsinization of the top layer of pig skin and maintained in serum-supplemented KBM medium.

For binding assays, we used MSSA ST398 isolates from human infections ($n = 2$; isolates ST398NM02 and MS298) and human-colonizing isolates ($n = 3$; isolates ST398NM01, 51368, and 51478). In addition, we used LA-ST398 MRSA isolates ($n = 5$; isolates M07-3, M07-11, M07-42, M07-43, and M07-98) recovered from healthy pigs in Ontario, Canada. All ST398 isolates were grown overnight in tryptic soy broth (TSB). Cultures were diluted 1/25, grown to mid-log phase, and adjusted to an optical density at 600 nm (OD_{600}) of 0.5, and 600 μl of bacteria were added to washed NHEK or porcine keratinocyte cell layers in triplicate. Samples were incubated for 1 h at 37°C, washed eight times with cell culture medium, and treated with 300 μl trypsin for 5 min at 37°C, and bacteria were lifted using a sterile cell scraper. Serial dilutions were plated on tryptic soy agar (TSA) plates with 5% sheep blood and were incubated overnight at 37°C. Alternatively, 10 μg/well of meta-keratin 4 (KeraFAST, Winston-Salem, NC) or recombinant cytokeratin-10 (32) (construct kindly provided by T. Foster, Dublin) were plated on 96-well plates and incubated overnight at 4°C. Wells were blocked with bovine serum albumin (BSA), and bacteria were added to triplicate wells and incubated for 1 h at 37°C. Samples were trypsinized and plated on TSA plates for overnight incubation. Results are representative of two independent experiments. Counts were compared using an independent *t* test (GraphPad; Prism).

Microarray data accession numbers. The array design is available in BμG@Sbase (accession no. A-BUGS-38; http://bugs.sgul.ac.uk/A-BUGS-38) and ArrayExpress (accession no. A-BUGS-38). Fully annotated microarray data have been deposited in BμG@Sbase (accession no. E-BUGS-124; http://bugs.sgul.ac.uk/E-BUGS-124) and also ArrayExpress (accession no. E-BUGS-124).

ACKNOWLEDGMENTS

This research was supported in part by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, NIH, and by NIH funding to A.-C.U. (K08 AI090013), F.D.L. (R01 AI077690 and R01 AI077690-S1), and W.I.L. (AI057158) and the Defense Threat Reduction Agency.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00027-12/-/DCSupplemental.

FIG S1, DOCX file, 0.1 MB.
FIG S2, DOCX file, 0.5 MB.
TABLE S1, DOCX file, 0.1 MB.
TABLE S2, DOCX file, 0.1 MB.
TABLE S3, DOCX file, 0.1 MB.

REFERENCES

1. **Lowy FD.** 1998. *Staphylococcus aureus* infections. N. Engl. J. Med. **339**:520–532.
2. **Leonard FC, Markey BK.** 2008. Meticillin-resistant *Staphylococcus aureus* in animals: a review. Vet. J. **175**:27–36.
3. **Devriese LA, Oeding P.** 1976. Characteristics of *Staphylococcus aureus* strains isolated from different animal species. Res. Vet. Sci. **21**:284–291.
4. **Hasman H, et al.** 2010. Spa type distribution in *Staphylococcus aureus* originating from pigs, cattle and poultry. Vet. Microbiol. **141**:326–331.
5. **Voss A, Loeffen F, Bakker J, Klaassen C, Wulf M.** 2005. Methicillin-resistant *Staphylococcus aureus* in pig farming. Emerg. Infect. Dis. **11**:1965–1966.
6. **Bens CC, Voss A, Klaassen CH.** 2006. Presence of a novel DNA methylation enzyme in methicillin-resistant *Staphylococcus aureus* isolates associated with pig farming leads to uninterpretable results in standard pulsed-field gel electrophoresis analysis. J. Clin. Microbiol. **44**:1875–1876.
7. **Stegger M, Lindsay JA, Sørum M, Gould KA, Skov R.** 2010. Genetic diversity in CC398 methicillin-resistant *Staphylococcus aureus* isolates of different geographical origin. Clin. Microbiol. Infect. **16**:1017–1019.
8. **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.
9. **van Cleef BA, et al.** 2011. Livestock-associated methicillin-resistant *Staphylococcus aureus* in humans, Europe. Emerg. Infect. Dis. **17**:502–505.
10. **Fan J, et al.** 2009. Biogeography and virulence of *Staphylococcus aureus*. PLoS One **4**:e6216.
11. **Khanna T, Friendship R, Dewey C, Weese JS.** 2008. Methicillin resistant *Staphylococcus aureus* colonization in pigs and pig farmers. Vet. Microbiol. **128**:298–303.
12. **van Loo I, et al.** 2007. Emergence of methicillin-resistant *Staphylococcus aureus* of animal origin in humans. Emerg. Infect. Dis. **13**:1834–1839.
13. **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.
14. **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.
15. **Cuny C, et al.** 2009. Nasal colonization of humans with methicillin-resistant *Staphylococcus aureus* (MRSA) CC398 with and without exposure to pigs. PLoS One **4**:e6800.
16. **Witte W, Strommenger B, Stanek C, Cuny C.** 2007. Methicillin-resistant *Staphylococcus aureus* ST398 in humans and animals, Central Europe. Emerg. Infect. Dis. **13**:255–258.
17. **Bootsma MC, Wassenberg MW, Trapman P, Bonten MJ.** 2011. The nosocomial transmission rate of animal-associated ST398 meticillin-resistant *Staphylococcus aureus*. J. R. Soc. Interface **8**:578–584.
18. **Bhat M, et al.** 2009. *Staphylococcus aureus* ST398, New York City and Dominican Republic. Emerg. Infect. Dis. **15**:285–287.
19. **Rasigade JP, Laurent F, Hubert P, Vandenesch F, Etienne J.** 2010. Lethal necrotizing pneumonia caused by an ST398 *Staphylococcus aureus* strain. Emerg. Infect. Dis. **16**:1330.
20. **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.
21. **Crombe F, et al.** 16 November 2011. Prevalence and antimicrobial susceptibility of methicillin-resistant *Staphylococcus aureus* among pigs in Belgium. Microb. Drug Resist. [Epub ahead of print.]
22. **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.
23. **Sung JM, Lloyd DH, Lindsay JA.** 2008. *Staphylococcus aureus* host specificity: comparative genomics of human versus animal isolates by multi-strain microarray. Microbiology **154**:1949–1959.
24. **Uhlemann AC, et al.** 26 July 2011. Molecular characterization of *Staphylococcus aureus* from outpatients in the Caribbean reveals the presence of pandemic clones. Eur. J. Clin. Microbiol. Infect. Dis. [Epub ahead of print.]
25. **McCarthy AJ, et al.** 2011. The distribution of mobile genetic elements (MGEs) in MRSA CC398 is associated with both host and country. Genome Biol. Evol. **3**:1164–1174.
26. **Uhlemann AC, et al.** 2011. The environment as an unrecognized reservoir for community-associated methicillin resistant *Staphylococcus aureus* USA300: a case-control study. PLoS One **6**:e22407.
27. **Tatusov RL, et al.** 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. **29**:22–28.
28. **Rocha EP, et al.** 2006. Comparisons of dN/dS are time dependent for closely related bacterial genomes. J. Theor. Biol. **239**:226–235.
29. **Castillo-Ramirez S, et al.** 2011. The impact of recombination on dN/dS within recently emerged bacterial clones. PLoS Pathog. **7**:e1002129.
30. **McCarthy AJ, Lindsay JA.** 2010. Genetic variation in *Staphylococcus aureus* surface and immune evasion genes is lineage associated: implications for vaccine design and host-pathogen interactions. BMC Microbiol. **10**:173.
31. **Walsh EJ, O'Brien LM, Liang X, Hook M, Foster TJ.** 2004. Clumping factor B, a fibrinogen-binding MSCRAMM (microbial surface components recognizing adhesive matrix molecules) adhesin of *Staphylococcus aureus*, also binds to the tail region of type I cytokeratin 10. J. Biol. Chem. **279**:50691–50699.
32. **Wassenberg MW, Bootsma MC, Troelstra A, Kluytmans JA, Bonten MJ.** 2011. Transmissibility of livestock-associated methicillin-resistant *Staphylococcus aureus* (ST398) in Dutch hospitals. Clin. Microbiol. Infect. **17**:316–319.

33. **Viana D, et al.** 2010. Adaptation of *Staphylococcus aureus* to ruminant and equine hosts involves SaPI-carried variants of von Willebrand factor-binding protein. Mol. Microbiol. **77**:1583–1594.
34. **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.
35. **Herron-Olson L, Fitzgerald JR, Musser JM, Kapur V.** 2007. Molecular correlates of host specialization in *Staphylococcus aureus*. PLoS One **2**:e1120.
36. **Phillips I, et al.** 2004. Does the use of antibiotics in food animals pose a risk to human health? A critical review of published data. J. Antimicrob. Chemother. **53**:28–52.
37. **DelVecchio VG, et al.** 2002. The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. Proc. Natl. Acad. Sci. U. S. A. **99**:443–448.
38. **Kapatral V, et al.** 2002. Genome sequence and analysis of the oral bacterium *Fusobacterium nucleatum* strain ATCC 25586. J. Bacteriol. **184**: 2005–2018.
39. **Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA.** 2011. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. BMC Genomics **12**:402.
40. **Kaneko J, Muramoto K, Kamio Y.** 1997. Gene of LukF-PV-like component of Panton-Valentine leukocidin in *Staphylococcus aureus* P83 is linked with lukM. Biosci. Biotechnol. Biochem. **61**:541–544.