Analysis of prophages harbored by the human-adapted subpopulation of Staphylococcus aureus CC398

Nathalie van der Mee-Marquet, Anna-Rita Corvaglia, Anne-Sophie Valentin, David Hernandez, Xavier Bertrand, Myriam Girard, Jan Kluytmans, Pierre-Yves Donnio, Roland Quentin, Patrice Francois

Analysis of prophages harbored by the human-adapted subpopulation of *Staphylococcus aureus* CC398

Nathalie van der Mee-Marquet[1,2*], Anna-Rita Corvaglia[3], Anne-Sophie Valentin[1], David Hernandez[3], Xavier Bertrand[4,5], Myriam Girard[3], Jan Kluytmans[6], Pierre-Yves Donnio[7,8], Roland Quentin[1], Patrice François[3]

[1]Service de Bactériologie et Hygiène, Centre Hospitalier Régional Universitaire, and [2]Réseau des Hygiénistes du Centre, Centre Hospitalier Universitaire, F37044 Tours cedex, France. [3]Genomic Research Laboratory, University of Geneva Hospitals, CH-1211 Geneva 14, Switzerland. [4]Service d'Hygiène, Centre Hospitalier Universitaire, and [5]UMR 6249 Chrono-environnement, Université de Franche-Comté, F25030 Besançon cedex, France. [6]Laboratory for Microbiology and infection Control, Amphia Hospital, Breda and VU University Medical Center, Amsterdam, The Netherlands. [7]Service de Bactériologie et Hygiène Hospitalière, Centre Hospitalier Universitaire, and [8]EA1254 Microbiologie-Risques Infectieux, Université de Rennes 1, F35033 Rennes cedex, France.

Short title: prophages in *Staphylococcus aureus* CC398

Corresponding author :
Nathalie van der Mee-Marquet
Laboratoire de Bactériologie et Hygiène
Hôpital Trousseau
37044 Tours cedex
France
Phone: 33 234 389 430
Fax : 33 247 478 588
email : n.vandermee@chu-tours.fr

ABSTRACT

Staphylococcus aureus clonal complex 398 is a livestock-associated pathogen that poses a worldwide threat because of its ability to colonize and infect both humans and animals. We used high-resolution whole-genome microarrays, prophage profiling, immune evasion cluster characterization and whole-genome sequencing to investigate the roles of prophages in the emerging human-adapted subpopulation of CC398 that has been associated with invasive infections in humans living in animal-free environments.

We characterized one phage and two prophages specifically harbored by CC398 isolates belonging to the emerging subpopulation. We introduced the phage into permissive prophage-free isolates. We investigated the effects of lysogeny on the host ability to resist further phage infection and transformation, to acquire the capacity to invade human cells, and to express virulence factors encoded by prophages. We report evidence of a defective ϕMR11-like helper prophage, named StauST398-5pro, specifically associated with the emerging non-LA CC398 subpopulation. StauST398-5pro confers substantial protection against horizontal genetic transfer to its host. It interacts with a human-associated β-converting prophage encoding immune-modulating proteins such that virulence genes are expressed during stress situations.

Our findings provide insight into the role of phages in the expression of virulence and in the spread of genetic information among new host-adapted *S. aureus* isolates. We demonstrate that functional prophage elements can condition host specificity and confer new virulence traits on emerging intra-species clones of bacteria.

Keywords: *Staphylococcus aureus*, sequence type 398, prophage, virulence, gene expression, host specificity, transformation, transduction.

INTRODUCTION

Originally detected in pig farmers in Europe (Witte, 2007), *S. aureus* belonging to CC398 has become a worldwide threat associated with livestock, their human contacts and food products (van Belkum, 2008). Microarrays and whole-genome sequencing approaches applied to a large number of CC398 isolates distinguished two clades within the CC398 lineage (Price, 2012; Uhleman, 2012; McCarthy, 2011): the classical LA clade, isolates of which have long been responsible for frequent and transient colonization, and rare infections, of farmers and veterinarians (Huijsdens, 2006); and a human clade. This human clade is comprised of two subpopulations: the ancestral human subpopulation, and the emerging human-adapted non-LA CC398 subpopulation that has recently and increasingly been causing invasive infections worldwide in humans living in animal-free environments (Price, 2012; Valentin-Domelier, 2011; Jimenez, 2011; Stegger, 2010), and that readily colonize and spread between humans (Uhleman, 2012).

Phages serve as a driving force in bacterial pathogenesis, contributing both to the evolution of bacterial hosts through gene transfer, and to bacterial pathogenesis at the time of infection (Wagner, 2002). *S. aureus* is highly lysogenic. *S. aureus* phages are double-stranded DNA phages belonging to the *Siphoviridae* family (Goerke, 2009). Comparative genomic analyses have revealed substantial diversity of these phages; nevertheless, they also display extensive mosaicism, with genes organized into functional modules that are frequently exchanged between phages (Kahankova, 2010; Xia, 2013). Temperate phages play an important role in the pathogenicity of *S. aureus*. Some carry genes encoding diverse virulence factors including panton-valentine leukocidin, staphylokinase, enterotoxins, chemotaxis-inhibitory proteins, and exfoliative toxins. Other phages have inserted into and

consequently interrupt chromosomal virulence genes such as those for β-hemolysin (*hlb*) and lipase (*geh*) (Goerke, 2009).

The two clades of the CC398 lineage are characterized by different prophages. LA CC398 isolates commonly carry phages ϕ2 and ϕ6 (Schijffelen, 2010; Hallin, 2011; McCarthy, 2011; Hallin, 2011), or a ϕAvβ prophage (Price, 2012). By contrast, isolates belonging to the human clade contain β-converting ϕ3 prophage variants that encode two immune-modulating proteins (Goerke, 2006; McCarthy, 2011; Price, 2012; Uhleman, 2012). These proteins are the staphylococcal complement inhibitor, SCIN, that prevents opsonophagocytosis and killing of *S. aureus* by human neutrophils; and the chemotaxis inhibitory protein of S. aureus, CHIPS. There is now evidence that the emerging subpopulation differs from the ancestral human LA subpopulation by additional prophage features (ϕ1, ϕ2, ϕ5 or ϕ7) relevant to its epidemiology (Uhleman, 2012; McCarthy, 2011). However, this issue has not been thoroughly investigated.

We aimed to characterize the prophage elements specifically associated with the *emerging* subpopulation. We applied whole genome analysis techniques to study the prophage content of isolates recovered in livestock environments and from cases of invasive infections in humans in animal-free environments. We studied the contribution of prophage content to phage resistance, competence for transformation and ability to invade human cells. We demonstrate that phages can condition host specificity and confer new virulence traits on pathogenic and host-adapted *S. aureus* strains.

MATERIALS AND METHODS

***Ethics Statement*.** The human isolates from The Netherlands were obtained during prospective surveillance, approved by the medical ethics committee of the Sint Elisabeth Hospital, at Tilburg, The Netherlands (protocol number 0749). All participants provided written informed consent. The strains isolated from animals were obtained by non-invasive sampling during a field study. This study was performed in accordance with the Dutch guidelines and as such did not require the approval of an ethics committee. The farmers gave informed consent and agreed to the collection of samples. The French isolates were obtained from clinical samples obtained during annual surveillance studies that were run according to the French Healthcare recommendations for prevention of infection. Ethical approval of the surveillance programs was obtained at the national level from the Réseau Alerte Investigation Surveillance des Infections Nosocomiales (RAISIN). In accordance with national French legislation and following the national procedure established by RAISIN, the surveillance study was run jointly with the regional surveillance coordinator, the director of the participating healthcare institutions and the physicians responsible for caring for the patients. The director and the physicians provided written consent for participation in the study. All patients or their relatives were individually approached for oral consent for participation in the study.

***Bacterial isolates*.** Twenty-one CC398 isolates were studied (Tab. 1), comprising two populations of isolates: (i) ten isolates representative of the diversity of the LA clade, including nine Dutch and one French isolates, recovered from humans (n=4) or animals (n=6), and belonging to the major *spa*-types t899 (n=2), t011 (n=2), t034 (n=2), t571 (n=2) and t108 (n=2); (ii) eleven isolates recovered over a four-year period in three geographically separated French regions (Valentin-Domelier, 2011) from patients with no LA risk factors

and diagnosed with bloodstream infections (n=10) or urinary tract infection (n=1). These isolates belonged to six different *spa*-types: t899 (n=1), t1451 (n=2), t571 (n=5), t6605 (n=1), t5635 (n=1) and t9378 (n=1).

DNA microarray experiments. The microarray has been described previously (Charbonnier, 2005). Microarrays were produced by *in situ* synthesis of a set of 15,600 60-mer long oligonucleotide probes (Agilent, Palo Alto, CA, USA). The probe set included 8,877 probes covering approximately 95 % of all ORFs annotated in strains N315 and Mu50 (Kuroda, 2001), MW2 (Baba, 2002), COL (Gill, 2005), NCTC 8325 (Gillaspy, 2006), USA300 (Diep, 2006), MRSA252 and MSSA476 (Holden, 2004) and including those on their plasmids. Each gene was covered by one to 12 probes depending on gene length. Each test DNA (test channel) was hybridized against the reference DNA (a mixture of DNA from all four strains used to design the array) in a single hybridization reaction. Fluorescence intensities were extracted using Feature Extraction software (Agilent, version 6.1.1). The signal was analyzed with in-house software that estimates the presence probability (EPP) for the sequence of each oligonucleotide probe, as previously described (Charbonnier, 2005). EPP values ≤1% were extracted and considered to indicate the absence of the sequence.

Prophage profiling and immune evasion cluster (IEC) characterization. A PCR-based assay for 47 prophage genes belonging to nine main phage modules was used (Kahankova, 2010). To characterize the IEC of the prophages, PCR was used to test for sequences corresponding to the prophage genes *hlb*, *scn*, *chp*, *sak*, *sep* and *sea* by the method developed by van Wamel (van Wamel, 2006). The primers described in Suppl. File Tab. 2 were used to test genomic and phage lysate DNA for *chp* and *scn*.

***Isolation, propagation and characterization of phages*.** The CC398 isolates were each assessed for their potential as a donor strain. Cultures were treated with mitomycin C

as described previously (de Gialluly, 2003). The filtered supernatants, corresponding to putative lysates, were stored at −80°C. The lytic activity of the putative lysates was tested against the 27 strains of the international phage typing system (de Gialluly, 2003) and the 21 CC398 isolates used as indicator strains. A putative lysate was considered to be a phage suspension if it provided at least 20 clear plaques on at least two indicator strains. Phages were propagated using the non-LA isolate S100. Titrated phage preparations were kept at −80°C. The induced phage preparations were centrifuged on a sucrose density gradient. Phage particles were negatively stained with 2% uranyl acetate, examined in a JEOL 1230 transmission electron microscope at an accelerating voltage of 120 kV, and photographed. Morphological types were defined on the basis of phage tail length.

Whole genome sequencing. The genomes of CC398 isolates S100, S124, S124/S100-ϕ were sequenced as previously described (Gizard, 2013), and the prophage contents assessed by searching for conserved phage genes and known phage-encoded virulence genes. The sequences are available through GenBank resources (see Tab. 2).

Study of the effects of lysogeny on resistance to phages, competence for transformation and invasion of human cells. The two prophage-free isolates (S123 and S124) were exposed to the lysate produced by the non-LA isolate S100. Cells surviving after incubation at 30°C for 18 hours were plated on fresh agar plates, incubated at 37°C for 18 hours and screened for their prophage content by PCR as described above. A prophage-containing isolate (transductant) was isolated. ***Phage resistance.*** To determine susceptibility to *S. aureus* phages, the 21 CC398 isolates were phage-typed using the international set of phages according to the procedure described previously (de Gialluly, 2003). ***Transformation experiments.*** DNA of pCN38 was prepared from *E. coli* DH5α and *Staphylococcus* RN4220 (Lee, 1995; Gizard, 2013). DNA of pMW401 was prepared from *Enterococcus faecalis* as

169 previously described (Beaume, 2010). The media used for overnight culture and subsequent
170 mini-preps (Qiaprep Spin Miniprep Kit, Qiagen) included Luria Bertani Broth (LB Broth,
171 Miller, BD Difco) for *E. coli* DH5α, supplemented with 100 mg/L of ampicillin; Mueller-Hinton
172 broth (MHB; BD Difco) for *S. aureus* RN4220; and Tryptic soy Broth for *E. faecalis* (Soyban-
173 Casein Digest Medium, BD Bacto). Chloramphenicol (10 mg/L) was included as appropriate.
174 Preparation of competent cells and transformation of RN4220, S100, S123, S124, S124/S100-
175 ϕ, S92, S93 and S94 were as described previously (Beaume, 2010). Following
176 electroporation, transformants were plated on Mueller-Hinton Agar, supplemented with
177 chloramphenicol: 10 mg/L for RN4220; 15 mg/L for S100, S123, S124, S124/S100-ϕ; and 20
178 mg/L for S92, S93, S94. The number of transformants obtained per µg of plasmid DNA was
179 calculated. ***Internalization experiments.*** The Cowan strain is internalized by human cells and
180 was therefore used as a positive control. Strain KH11 was used as a non-invasive, negative
181 control. The internalization procedure and the lysostaphin protection assay were based on a
182 published method (Sinha, 1999) with minor differences. Briefly, cultures of the cell line 293
183 were maintained under humidified air with 5 % CO_2 at 37°C in Dulbecco's minimal essential
184 medium (DMEM)-nutrient mixture F-12 (nut mix F-12) containing Glutamax I, a stable
185 glutamine dipeptide, supplemented with 10 % fetal calf serum (FCS), 50 IU of penicillin per
186 mL, and 50 µg of streptomycin per mL. For the lysostaphin assay, 293 cells were plated in
187 poly-ornithine coated plates (1.8 cm diameter) and cultured for 24 hours, such that there
188 were approximately 0.9×10^5 cells/well. Bacterial were harvested from culture, washed,
189 counted and diluted (Sinha, 1999); these bacterial dilutions were added to the plates
190 containing 293 cells at a multiplicity of infection [MOI] of 15:1. The plates were incubated for
191 1 h at 37°C, and the bacterial suspensions were then replaced with lysostaphin medium
192 (DMEM-nut mix F-12, 10 % FCS, 20 µg of lysostaphin per mL). The plates were further

incubated for 30 min, and the 293 cells then lysed in 1 mL of sterile distilled H_2O. Appropriate serial dilutions of the cell lysate in PBS were plated on Mueller-Hinton agar or TSA plates, and CFU were counted manually. ***Transcription experiments.*** Gene-specific primers and probes were designed using Primer Express 3.0 (Applied Biosystems) and are shown in Suppl. File Tab. 2. Oligonucleotide primers and probes were used at final concentrations of 0.2 and 0.1 µM, respectively, in a final volume of 10 µL of one-step RT–PCR enzymatic mixture (Invitrogen, Carlsbad, Germany); a Mx3005P system (Agilent) was used as described previously (Beaume, 2010). Results were normalized to those for the *hu* gene, as described previously. These assays provided relative gene expression values for *chp*, *scn* and *sak* in various isolates and in various stress conditions.

Statistical data. Chi-square tests and Fisher's exact test (two-tailed) were used to test associations; a P value <0.05 was considered significant.

205 **RESULTS**

206 **1. IDENTIFICATION OF CC398 ISOLATES BELONGING TO THE EMERGING SUBPOPULATION.**

207 We studied the prophage content of 21 CC398 isolates to identify those belonging to
208 this *emerging* subpopulation. We searched for isolates carrying a β-converting prophage
209 encoding an immune evasion cluster (IEC) —characteristic of the two subpopulations of the
210 human clade (Price, 2012; Uhleman, 2012; McCarthy, 2011)— and additional prophages
211 elements, believed to distinguish isolates of the ancestral subpopulation from the emerging
212 human subpopulation (Price, 2012). Using three different methods, IEC characterization,
213 prophage profiling and hybridization of genomic DNA with microarrays, the ten non-LA
214 MSSA isolates were all distinguished from the LA clade and confirmed to belong to the
215 human clade**.** An IEC carrying *chp* and *scn* (van Wamel, 2006) was found in all non-LA
216 isolates, and no IEC was found in any of the LA isolates (Tab. 1). LA lysogens showed a
217 diversity of prophage profiles, all of which lacked a Sa3 integrase; and their DNA hybridized
218 with probes representing the L54a phage (Fig. 1). In contrast, the prophage profiles of non-
219 LA isolates were more homogeneous: all prophages in these isolates had genes encoding a
220 Sa3 integrase and F serogroup; and the DNA from all the non-LA isolates hybridized with
221 probes corresponding to a β-converting prophage encoding the putative virulence factor
222 CHIPS. This indicates that all these non-LA isolates contain a β-converting prophage encoding
223 an IEC.

224 Prophage profiling and microarray data revealed two subgroups among the non-LA
225 MSSA isolates (Tab.1; Fig.1). One subgroup comprised six isolates only harboring Sa3
226 prophage elements. Four isolates in the second subgroup harbored additional prophage
227 elements next to the Sa3 prophage elements. These additional elements hybridized with
228 probes for a gene encoding a protein related to SaPI, and probes for phages ϕSLT and ϕETA.

Some of these additional prophage elements were similar to those in the LA-isolates. Isolates belonging to the first of these two subgroups were considered to belong to the ancestral subpopulation, and those of the second subgroup to belong to the emerging subpopulation. These subgroups were named A, for ancestral subpopulation, and E, for emerging subpopulation respectively.

2. PHAGE ISOLATION.

For further characterization of prophage features specifically associated with the emerging subpopulation represented by isolates of subgroup E, we tested whether the prophage elements identified in their genomes encoded functional or defective prophages. The 21 isolates were treated with mitomycin C and culture filtrates obtained. Fourteen of the 21 filtrates, including those of subgroup E, produced plaques on indicator strains, indicating the presence of infectious particles (Suppl. file Tab. 1). LA-isolate cultures cleared, indicating cell lysis, and thus that the prophages induced were functional. In contrast, cell lysis was not observed in any non-LA isolate cultures following treatment with mitomycin C: even after prolonged incubation, these cultures remained turbid. This indicates that most non-LA cells were able to survive in the presence of their phage particles, suggesting that the lysogens had superinfection immunity of or that phages were defective. One infective lysate was studied further: S100-ϕ obtained from the non-LA isolate of subgroup E S100. The phage was purified and propagated. Electron microscopy analysis (Suppl. file Fig. 1) indicated that the phage belongs to the *Siphoviridae* family: phages of this family have isometric, non-enveloped heads and filamentous cross-banded tails with short terminal fibers. The phage particles in lysate S100-ϕ had 53nm-diameter heads and138nm-long tails (± 10nm).

3. Lysogenization of prophage-free isolates with the phages produced by S100.

A reliable method for studying the consequences of lysogeny is to lysogenize clonal recipient strains devoid of prophage elements with natural phage. Prophage profiling and whole-genome sequencing results demonstrated that the LA isolates S123 and S124 were devoid of prophages, and these isolates were therefore used as prophage-free recipients. The infective lysate S100-ϕ was used to transduce the prophage-free isolates S123 and S124. Lysogenization was only obtained with S124, leading to isolate S124/S100-ϕ. Lysogenization was confirmed by phage profiling and whole-genome sequencing of the transductants. S124/S100-ϕ lysogens were not stable, as the prophage was rapidly lost during bacterial culture.

4. Characterization of phages and prophages of the emerging human-adapted non-LA isolate S100.

Using whole-genome sequencing, we identified one phage, StauST398-1 in the non-LA S100-ϕ lysate. Whole-genome comparison of StauST398-1 with parental (S124), transductant (S124/S100-ϕ) and donor (S100) isolates revealed that the StauST398-1 phage resulted from the recombination of the prophages, StauST398-4pro and StauST398-5pro, in the S100 donor isolate genome (Fig. 2). This recombination has led to the loss of part of the IEC locus (Fig. 2).

The StauST398-4pro genome (accession number KC595278) is a linear, double-stranded DNA molecule of 41.9 Kb and 0.339% G+C. It contains 65 putative coding sequences and was inserted into the *hlb* gene, 824 nucleotides downstream from the ATG codon. Genes for lysogeny, including an integrase (RusA), a repressor protein, and an antirepressor, were found between genes involved in the lytic cycle and replication. No

277 known restriction system was found. However, part of a putative type III restriction-
278 modification system was identified by the annotating tool. The phage genome carries two
279 putative virulence genes, *chp* and *scn*, around 2.5 Kb upstream from the insertion site. The
280 sequence of StauST398-4pro shares approximately 97% identity with phi3, recently
281 described in *S. aureus* strain 71193 (Uhleman, 2012) and 91% with the close neighbor of
282 phiNM3, recently described by Price and colleagues (Price, 2012). The two other phages
283 described by Price and colleagues are much more divergent from StauST398-4pro (<15%
284 identity). PCR with primers specific for StauST398-4pro, amplified fragments of this β-
285 converting ϕ3-prophage from all of the non-LA CC398 isolates but none of the LA isolates
286 (Suppl. File Tab.2).

287 The StauST398-5pro genome (accession number KC595279) is a 40.2 Kb, linear,
288 double-stranded DNA with 0.353% G+C. It carries 55 putative coding sequences, including a
289 putative gene for a superantigen similar to enterotoxin B. It is inserted in *smpB*. It has a
290 modular organization with gene clusters involved in DNA packaging, head and tail
291 morphogenesis, cell lysis, lysogeny, replication and regulation. The lysogeny module was
292 found between the lysis cassette and the replication module, with a repressor protein cI, the
293 repressor cro and an antirepressor that appeared to result from antA/B recombination
294 (Davis, 2002). The transcriptional regulation module comprises a dimeric dUTPase, an
295 integrase activator and a rho terminator factor. This phage lacks an integrase, indicating a
296 lytic-defective prophage. Homology searches identified a putative restriction system.
297 StauST398-5pro shares 75% identity with ϕMR11 (Rashel, 2007) and with ϕETA (Yamaguchi,
298 2000) over more than half of its length. The remaining 15 Kb of the genome showed a highly
299 mosaic structure with small segments (generally <3 Kb) sharing some similarity with ϕPVL,

ϕETA, ϕ13 or ϕROSA (Yamaguchi, 2000; Kaneko, 1998; Iandolo, 2002). PCR tests with specific primers detected StauST398-5pro only in isolates of subgroup E (Suppl. File Tab.3).

The StauST398-1pro genome (accession number JX013863) is 45.2 Kb, 0.345% G+C, and contains 72 putative protein-coding genes. The 5-35Kb region of this phage genome is homologous to ϕETA (Yamaguchi, 2000) and has two 5 Kb regions of similarity with the genome extremities of ϕNM3 (Bae, 2006; Baba, 2008). The StauST398-1pro genome encodes a dUTPase and a putative endonuclease (*nuc*) also present in ϕNM3. The phage replicon is similar to *S. aureus* phage ϕ80 (Christie, 2010). It contains the IEC with only *scn* but not *chp,* which is 3 Kb upstream from the *attL* site in StauST398-4 pro. The recombinant phage StauST398-1 was longer than the prophage StauST398-1pro in the S124 genome, resulting from a partial deletion of the prophage during the insertion. To evaluate the integrity of the IEC during phage excision and insertion, phage lysate S100-ϕ and genomic DNA from isolates S100, S124 and S124/S100-ϕ were subjected to PCR to test for the presence of *chp* and *scn*. Both genes were present in StauST398-1, whereas only *scn* was detected in StauST398-1pro. This indicates that *chp* was lost upon integration of StauST398-1 into the S124 genome.

5. IMPACT OF LYSOGENY ON PHAGE RESISTANCE, TRANSFORMATION AND INVASION OF HUMAN CELLS.

To investigate whether lysogeny in CC398 isolates of subgroup E affected various host characteristics, we tested lysogenic and non-lysogenic *S. aureus* for (i) resistance to infection by a large panel of *S. aureus* phages, (ii) transformation by vectors propagated in *S. aureus* or other hosts, and (iii) epithelial cell invasiveness.

5.1. Phage resistance. The prophage-free isolates S123 and S124 were highly susceptible to phage infections; by contrast, all LA *and* non-LA lysogens were resistant to

numerous, diverse *S. aureus* phages (Suppl. file Tab. 1). Lysogenization of the prophage-free isolate S124 with S100-ϕ lysate, containing the recombinant phage StauST398-1, resulted in resistance to a broad range of phages, similar to isolates lysogenized with the ϕ42 prophage that encodes a powerful restriction modification (RM) system (Dempsey, 2005).

5.2. Transformation experiments*.* The efficiency with which the following strains could be transformed was determined: the parental strains (S123 and S124), transductant strain (S124/S100-ϕ), non-LA donor S100 of subgroup E (harboring StauST398-4pro and StauST398-5pro) strains, three non-LA isolates of subgroup A (that only harbor StauST398-4pro), and *S. aureus* RN4220, a laboratory strain lacking a RM system (Nair, 2011) (Tab. 2). The prophage-free isolates S123 and S124 were transformed efficiently by both RN4220 and *E. coli* plasmid DNA. Plasmid DNA from *Enterococcus faecalis*, a potential source of the *van* operon, was also taken up by these two isolates at a frequency similar to that for *S. aureus* RN4220. By contrast, the naturally lysogenic isolates S1 and S100 were resistant to transformation with these plasmid DNAs. The efficiency of transformation of the S124 prophage-free isolate was substantially reduced by lysogenization with the non-LA infective lysate. Note that the reduction of transformation efficiency upon lysogenization may have been partially masked because of the instability of lysogeny observed in S124/S100-ϕ. The isolates of subgroup A, that harbor the IEC-containing prophage StauST398-4pro and lack StauST398-5pro, were efficiently transformed with plasmid DNA (RN4220 or *E. coli* DH5); this indicates a weaker protection against transformation by a broad range of foreign DNAs than observed with the isolates of subgroup E that harbor the two prophages (StauST398-4pro and StauST398-5pro).

5.3. Invasion experiments*. S. aureus* is not an obligatory intracellular pathogen. However, the intrinsic virulence of isolates is associated with their ability to reside in the

cytoplasm of host epithelial or endothelial cells, particularly in chronic infections (Sinha, 1999; Voyich, 2005). Invasion of epithelial cells by representative isolates (the parental (S123 and S124) and donor (S100) strains, LA isolates (S1 and S130), non-LA isolates (S89 and S94), and two control strains (Cowan and KH11) was evaluated. Note that the transductant S124/S100-ϕ could not be studied, because of rapid loss of the phage during the experiments. All the LA and non-LA lysogens were 3-4 fold more invasive than the Cowan strain (Fig. 3), a highly invasive strain used as a positive control. The prophage-free isolates S123 and S124 were not able to invade cells; however, lysogenization conferred invasiveness on these strains.

Transcription experiments. RNA was isolated during exponential growth and stationary phase to test for the expression of virulence-related genes of the IEC locus (*chp, sak* and *scn*). The following non-LA isolates were studied: S81 carrying StauST398-4pro and StauST398-5pro (subgroup E), S94 and S89 only carrying StauST398-4pro. The unstable transductant S124/S100-ϕ was not studied. In all three isolates, the expression of virulence-related genes was strongly induced during stress, particularly exposure to concentrations of antibiotics below the minimum inhibitory concentration (MIC) (Fig. 4). This is consistent with the general increase in gene expression during stress, particularly during infection (Beaume, 2010). Virulence-related mRNAs were similarly abundant during stationary and exponential growth phases in both S94 and S89, two isolates that only harbour StauST398-4pro; however, these mRNAs were more abundant in S81, the isolate carrying StauST398-4pro and StauST398-5pro, during the stationary phase than the exponential growth phase.

DISCUSSION

Using a novel approach, different to those used previously (Price, 2012; Uhleman, 2012; McCarthy, 2011), we report prophage features specifically associated with the CC398 isolates that have recently been causing an increasing number of invasive infections worldwide in humans living in animal-free environments (Stegger, 2010; Valentin-Domelier, 2011; Jimenez, 2011) and that readily colonize humans and spread between them (Uhleman, 2012). We used a variety of methods, including genomic microarray analysis, phage content profiling, immune-evasion cluster typing and genome sequencing, to detect and analyze these genetic elements. We report evidence for a helper prophage, not previously described, that interacts with the well-recognized human-associated β-converting prophage, encodes a restriction modification system, and influences the ability of CC398 isolates to colonize and infect humans.

We used genotypic methods allowing subpopulations to be distinguished within an apparently homogeneous population of *S. aureus* isolates. It has limitations. One is the way we used of microarrays, which allowed analysis of only known or identified target sequences. Also the choice of strains included in the analysis may have affected the findings, and the number of strains studied was limited. We are currently investigated a larger collection of strains. We were nevertheless able to identify particular phenotypes specific to groups of strains and at least partly relate these phenotypes to the genome content of representative isolates used for full genome sequencing.

The characterization of prophages from naturally lysogenic non-LA isolates provided evidence of two distinct subgroups among MSSA isolates responsible for human invasive infections. The first subgroup includes isolates that only carry a Sa3 prophage (named StauST398-4pro). The second subgroup comprises isolates carrying this same Sa3 prophage,

393 but also a ϕMR11-like prophage (named StauST398-5pro). StauST398-4pro is an IEC β-
394 converting Sa3 prophage that has been described in isolates responsible for human
395 infections in LA- and non-LA environments (Stegger, 2010; Valentin-Domelier, 2011;
396 Jimenez, 2011), but StauST398-5pro has not previously been detected in isolates responsible
397 for human infections in a LA environment. Therefore, we suggest that isolates only harboring
398 StauST398-4pro belong the ancestral human LA subpopulation, and that StauST398-5pro is
399 specific to isolates responsible for the more recent human invasive infections in a non-LA
400 environment, i.e. isolates of the emerging human-adapted non-LA CC398 subpopulation.

401 Lysogeny is a symbiotic strategy in which the prophage has a stable niche in the host
402 genome, and protects its host against attack by other phages (Goerke, 2009), through
403 co-immunity, CRISPR interference (Barrangou, 2007; Marraffini, 2008) or expression of RM
404 activities that inhibit interspecies and even intraspecies gene transfers (Barrangou, 2007;
405 Park, 1999). This may be particularly relevant for the CC398 lineage. Our sequencing findings
406 clearly indicate that CC398 does not harbor any genes for a type IV RM system similar to
407 those previously described (Corvaglia, 2010). In addition, the 296 bp fragment of the *sauI-*
408 *hsdSI* gene, a reliable discriminative marker of *S. aureus* CC398 (Stegger, 2011), is present
409 and strictly conserved in our isolates. However, our transformation and transduction
410 experiments clearly indicate that the presence of this *sauI-hsdSI* gene, a component of a
411 type I RM system, is not associated with protection against uptake of foreign DNA (Waldron,
412 2006; Corvaglia, 2010).

413 We demonstrate that CC398 prophages, like most described *S. aureus* prophages,
414 enabled their hosts to exclude other phages and exogenous DNA. However, the protection
415 conferred by these prophages varied: the protection conferred by StauST398-4pro prophage
416 was however limited, but if StauST398-5pro prophage was present in the non-LA isolate, the

417 protection was strong. *In silico* analysis of complete genomes from naturally lysogenic non-
418 LA CC398 did not identify any CRISPR or CRISPR-associated genes. Prophage sequences
419 encoding a putative RM system were identified in StauST398-5pro. Thus, the emerging
420 CC398 isolates, all of which harbor StauST398-5pro, have a prophage-encoded RM system
421 conferring strong protection against horizontal genetic transfer, similar to that described for
422 the ϕ42 prophage (Dempsey, 2005). This protection may confer a selective advantage over
423 competing organisms in their environment.

424 When prophages are fully functional, lysogeny is an Achilles heel for the bacteria;
425 they are vulnerable to environmental factors that cause sublethal DNA damage and
426 subsequent prophage induction. Prophage repressors are inactivated during the SOS
427 response, allowing the phage lytic cycle to proceed (Davis, 2002; Ubeda, 2005). Non-LA
428 isolates were able to produce phage particles, but only following recombination between
429 two defective prophages: the IEC (*chp-scn*) β-converting *S. aureus* prophage variant that has
430 previously been reported to be defective (van Wamel, 2006), and StauST398-5pro, the
431 genome of which does not contain an integrase gene. If integrated prophages become
432 defective, SOS-induced prophages do not lyse cells such that the host cells become
433 insensitive to the bactericidal effects of DNA-damaging antibiotics. This situation is thus
434 similar to that of cells that are not lysogens. We observed this phenomenon during the non-
435 LA phage lytic cycle. Therefore, in the human environment, where the SOS response is
436 frequently activated (Ubeda, 2005), natural lysogens belonging to the *emerging*
437 subpopulation and harboring the defective prophages StauST398-4pro and StauST398-5pro
438 may not be less vulnerable.

439 Prophages can contribute directly to bacterial pathogenesis at the time of infection
440 (Sumby, 2003). The expression of temperate prophage genes, that are generally repressed,

441 may be induced during stress and/or induction of the SOS system, because the repressor, cI,
442 is inactivated by RecA. This can result in increased transcription of virulence factors, as has
443 been demonstrated for staphylokinase production in *S. aureus* strains carrying β-converting
444 phages with the *sak* gene (Rooijakkers, 2006). We demonstrated that the non-LA prophage
445 StauST398-5pro encodes a putative superantigen similar to enterotoxin B and the non-LA β-
446 converting prophage StauST398-4pro encodes and expresses two human-specific virulence
447 genes (*chp* and *scn*) (van Wamel, 2006). *S. aureus* enters the body through breaches in the
448 skin or mucous membranes, and is then immediately confronted by the host innate defense
449 system. To counteract innate immunity, *S. aureus* expresses immune-modulating proteins
450 encoded by β-converting prophages, that facilitate long-term colonization in humans
451 (Goerke, 2006; Rooijakkers, 2006): the staphylococcal complement inhibitor (SCIN) prevents
452 opsonophagocytosis and killing of *S. aureus* by human neutrophils, the chemotaxis inhibitory
453 protein of *S. aureus* (CHIPS), the bacterial plasminogen activator (SAK) and the superantigens
454 SEA and SEB, that down-regulate chemokine receptors on monocytes. Our transcription
455 experiments indicate that non-LA CC398 lysogens may benefit from the production of SEB,
456 CHIPS and SCIN (products of *seb*, *chp* and *scn* genes, respectively) when exposed to
457 conditions that favor prophage induction.

458 In the genome of StauST398-5pro, we identified a transcription repressor and
459 antirepressor resulting from antA/B recombination, characteristic of a helper prophage
460 (Davis, 2002). The association between a helper-phage and a satellite prophage allows the
461 production proteins encoded by the satellite phage that may escape lysogenic repression
462 because of interactions between repressors and antirepressors. In *S. aureus*, the association
463 between a β-converting phage and a helper phage has been shown to result in the
464 expression of prophage-encoded virulence genes, such as *sea* and *sak*, during stress (e.g.

DNA damage) and lysogeny (Coleman, 1989; Carroll, 1993). Therefore, lysogeny associating the ϕMR11-like helper prophage StauST398-5pro (encoding *seb*) and the satellite β-converting prophage StauST398-4pro (encoding *chp* and *scn*) may facilitate the expression of prophage virulence genes, notably *chp* and *scn*, both when the prophage is induced and during lysogeny.

Lysogenization of *S. aureus* with CC398-phages conferred increased invasiveness. It seems unlikely that phage-encoded lysins affect human epithelial cells, as their target structures are not found in eukaryotic cells. However, phage endopeptidases, for example lysostaphin which is capable of degrading some human proteins, may affect mammalian tissues (Park, 1999). This possibility could be investigated with an animal model of endocarditis for example, as CC398 isolates have been associated with a human case of endocarditis in a highly immuno-compromised patient (Schijffelen, 2010). However, the two subgroups of isolates recovered from human invasive infections (those harboring StauST398-4pro and those harboring StauST398-4pro and StauST398-5pro) did not differ in their ability to invade human cells, suggesting that the StauST3985-pro prophage is not involved.

The prophages StauST398-4pro and StauST398-5pro were inserted at different loci in the genomes of non-LA isolates, but it all cases inactivated a bacterial virulence gene: StauST398-4pro in *hlb* and StauST398-5pro in *smpB.* Phages are potentially unstable and bacterial isolates tend to loose their phages. Thus, the maintenance of a β-converting prophage in the host chromosome indicates that this phage is beneficial or even essential. Such prophages may complement strain physiology and allow survival in the potentially hostile environmental conditions during colonization and infection (Smeltzer, 1994; Wagner, 2002). Studying *S. aureus* and diabetic feet, Lavigne recently associated the disruption of a staphylococcal virulence gene by a ROSA-like prophage with the colonizing status of

S. aureus isolates (Messad, 2012). The authors demonstrated that lysogeny resulted in the human immune response being limited thereby favoring resilient and stable colonization. StauST398-4pro and StauST398-5pro in CC398 isolates may have a similar effect; virulence may be reduced by prophage integration, favoring colonization, and thus bloodstream infection (del Rio, 2009).

Concordant with recent studies, our findings support the existence of three subpopulations in the CC398 lineage: an ancestral clade that originated in humans (here represented by subgroup A), a lineage composed predominantly of livestock-associated isolates (Price, 2012), and one that emerged recently by clonal expansion in humans (here, subgroup E) (Uhleman, 2012). Indeed, hybridization of genomic DNA with microarrays revealed no LA-prophage remnants in the genomes of the first subgroup of non-LA isolates, consistent with this subgroup being part of the human ancestral clade, and not the classical livestock-associated clade. In contrast, our data demonstrate that subgroup E of non-LA isolates and the LA isolates share some prophage elements, suggesting an animal origin of the newly described prophage StauST398-5pro. Further investigations are required to confirm these findings.

CONCLUSION

By studying genomes using microarrays, and determining prophage content and lysogeny characteristics within the CC398 lineage, we describe important features of the human-adapted emerging non-LA CC398 subpopulation. The association of a pre-existing β-converting-defective phage and a newly infecting phage possessing the characteristics of a helper phage, has resulted in pathogen diversification via gene transfer, and contributes to bacterial pathogenesis at the time of infection. We clearly demonstrate that prophage

513 content can influence the phenotypes that contribute to genome plasticity (notably
514 transformation and transduction) and to the invasiveness of *S. aureus*. These prophages in
515 the bacterial host provide protection against horizontal genetic transfer, and also contribute
516 to virulence and adaptation of the host bacterium.
517

ACKNOWLEDGMENTS

We thank Professor Alex van Belkum for comments and useful suggestions. The members of the Bloodstream Infection Study Group of the Réseau des Hygiénistes du Centre are AMIRAULT P (Vierzon), ARCHAMBAULT M (Pithiviers), BACHELIER MN (Bourges), BLOC D (Tours), BOUCHER M (Chateaudun), CATTIER B (Amboise), CHANDESRIS C (Amilly Montargis), CHEVEREAU V (La Chaussée St Victor), COUROUBLE G (Chateauroux), COURTIN M.-C (Amboise), DECREUX C (Chateauroux), DE GIALLULY C (Tours), DENIS C (Loches), DEPERROIS F (Chinon), FIEVRE C (Le Blanc), FOLOPPE P (Loches), FONGAUFFIER F (Chateaudun), FOURNIER-HOOCK R (Amilly Montargis), GIRARD N (Tours), GOURDET T (La Chaussée St Victor), GRAVERON JL (Fleury Les Aubrais), GROBOST F (La Ferté Bernard), GUILLON MF (Chateauroux), GUINARD F (Bourges), HARRIAU P (St Amand Montrond), HOMBROUCK-ALET C (Blois, Vendome, Romorantin), IMBAULT D (Vendome), JEHANNO D (Fleury Les Aubrais), KOURTA MJ (Chateaudun), LAURENT O (St Doulchard), LEHIANI O (Vierzon, Bourges, St Amand Montrond), LEPINEUX DA ROCHA A (St Amand Montrond), LESIMPLE AL (Vendome), LOUVIER X (Gien), MICHEL V (Le Blanc), MORANGE V (Tours), MOREL-DESJARDINS E (Bourges), MORIN E (Orléans), NAUDION C (Romorantin), NARBEY D (Blois), NEVEU C (Dreux), PABA O (Vendome), PERIGOIS F (Le Blanc), PETIT LE GOUAS G (Nogent Le Rotrou), POITVIN D (Chinon), PREVOST-OUSSAR M (Pithiviers), RATOVOHERY D (Chateauroux), ROUSSEAU B (Gien), ROUSSIN A (Orléans), SECHER A (Dreux). WATT S (Chinon).

REFERENCES

- Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K. et al., 2002. Genome and virulence determinants of high virulence community-acquired MRSA. Lancet. 359, 1819-27.
- Baba, T., Bae, T., Schneewind, O., Takeuchi, F., Hiramatsu, K., 2008. Genome sequence of *Staphylococcus aureus* strain Newman and comparative analysis of staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands. J. Bacteriol. 190, 300-10.
- Bae, T., Baba, T., Hiramatsu, K., Schneewing, O., 2006. Prophages of *Staphylococcus aureus* Newman and their contribution to virulence. Mol. Microbiol. 62, 1035-47.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P. et al.,2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 315, 1709-12.
- Beaume, M., Hernandez, D., Farinelli, L., Deluen, C., Linder, P. et al.,2010. Cartography of methicillin-resistant *Staphylococcus aureus* transcripts: detection, orientation and temporal expression during growth phase and stress conditions. PLoS One. 5, e10725.
- Carroll, J.D., Cafferkey, M.T., Coleman, D.C., 1993. Serotype F double- and triple-converting phage insertionally inactivate the *Staphylococcus aureus* beta-toxin determinant by a common molecular mechanism. FEMS. Microbiol. Lett. 106, 147-55.
- Charbonnier, Y., Gettler, B., Francois, P., Bento, M., Renzoni, A. et al., 2005. A generic approach for the design of whole-genome oligoarrays, validated for genomotyping, deletion mapping and gene expression analysis on *Staphylococcus aureus*. BMC. Genom. 6, p. 95.

- Christie, G.E., Matthews, A.M., King, D.G., Lane, K.D., Olivarez, N.P. et al., 2010. The complete genomes of *Staphylococcus aureus* bacteriophages 80 and 80α--implications for the specificity of SaPI mobilization. Virology. 407, 381-90.
- Coleman, D.C., Sullivan, D.J., Russell, R.J., Arbuthnott, J.P., Carey, B.F., Pomeroy, H.M., 1989. *Staphylococcus aureus* bacteriophages mediating lysogenic conversion of beta-lysin, staphylokinase and enterotoxin A: molecular mechanism of triphe conversion. J. Gen. Microbiol. 135, 1679-97.
- Corvaglia, A.R., François, P., Hernandez, D., Perron, K., Linder, P. et al., 2010. A type III-like restriction endonuclease functions as a major barrier to horizontal gene transfer in clinical *Staphylococcus aureus* strains. Proc. Natl. Acad. Sci. U S A. 107, 11954-58.
- Davis, B.M., Kimsey, H.H., Kane, A.V., Waldor, M.K., 2002. A satellite phage-encoded antirepressor induces repressor aggregation and cholera toxin gene transfer. EMBO. J. 21, 4240-49.
- de Gialluly, C., Loulergue, J., Bruant, G., Mereghetti, L., Massuard, S. et al. ,2003. Identification of new phages to type *Staphylococcus aureus* strains and comparison with a genotypic method. J. Hosp. Infect. 55, 61-7.
- del Rio, A., Cervera, C., Moreno, A., Moreillon, P., Miró, J.M. 2009. Patients at risk of complications of *Staphylococcus aureus* bloodstream infection. Clin. Infect. Dis. 48 Suppl 4, S246-53.
- Dempsey, R.M., Carroll, D., Kong, H., Higgins, L., Keane, C.T. et al.. 2005. Sau42I, a BcgI-like restriction-modification system encoded by the *Staphylococcus aureus* quadruple-converting phage Φ42. Microbiol. 151, 1301-11.

- Diep, B.B., Gill, S.R., Chang, R.F., Phan, T.H., Chen, J.H. et al. 2006. Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant *Staphylococcus aureus*. Lancet. 367,731-9.
- Gill, S.R., Fouts, D.E., Archer, G.L., Mongodin, E.F., Deboy, R.T., Ravel, J. et al., 2005. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. J. Bacteriol. 187, 2426-38.
- Gillaspy, A.F., Worrell, V., Orvis, J., Roe, B.A., Dyer, D.W. et al. 2006. Gram-positive pathogens. 2nd edition. Edited by Fischetti, V.A., Novick, R.P., Ferretti, F., Portnoy, D.A., Rood, J. Washington, DC, ASM Press. pp. 381-412.
- Gizard, Y., Zbinden, A., Schrenzel, J., François, P. 2013. Whole-Genome Sequences of *Streptococcus tigurinus* Type strain AZ 3a and *S. tigurinus*, a strain causing prosthetic joint infection. Genome Announc. 1(2). doi:pii: e00210-12. 10.1128/genomeA.00210-12.
- Goerke, C., Kraning, K., Stern, M., Döring, G., Botzenhart, K. et al., 2004. Increased frequency of genomic alterations in *Staphylococcus aureus* during chronic infection is in part due to phage mobilization. J. Infect. Dis. 189, 724-34.
- Goerke, C., Wirtz, C., Flückiger, U., Wolz, C., 2006. Extensive phage dynamics in *Staphylococcus aureus* contributes to adaptation to the human host during infection. Mol. Microbiol. 61, 1673-85.
- Goerke, C., Pantucek, R., Holtfreter, S., Schulte, B., Zink, M. et al., 2009. Diversity of prophages in dominant *Staphylococcus aureus* clonal lineages. J. Bacteriol. 191, 3462-68.
- Hallin, M., De Mendonca, R., Denis, O., Lefort, A., El Garch, F. et al. 2011. Diversity of accessory genome of human and livestock-associated ST398 methicillin resistant *Staphylococcus aureus* strains. Infcet. Genet. Evol. 11,290-9.

607 - Huijsdens, X.W., van Dijke, B.J., Spalburg, E., van Santen-Verheuvel, M.G., Heck, M.E. et
608 al., 2006. Community-acquired MRSA and pig-farming. Ann. Clin. Microbiol. Antimicrob.
609 5, 26.
610 - Holden, M.T., Feil, E.J., Lindsay, J.A., Peacock, S.J., Day, N.P. et al. 2004. Complete
611 genomes of two clinical Staphylococcus aureus strains: evidence for rapide evolution of
612 virulence and drug resistance. Proc. Natl. Acad. Sci. U S A. 101,9786-91.
613 - Iandolo, J.J., Worrell, V., Groicher, K.H., Qian, Y., Tian, R., Kenton, S. et al., 2002.
614 Comparative analysis of the genomes of the temperate bacteriophages ϕ11, ϕ12 and
615 ϕ13 of *Staphylococcus aureus* 8325. Gene. 289, 109-18.
616 - Jimenez, J.N., Velez, L.A., Mediavilla, J.R., Ocampo, A.M., Vanegas, J.M. et al., 2011.
617 Livestock-associated methicillin-susceptible *Staphylococcus aureus* ST398 infection in
618 women, Colombia. Emerg. Infect. Dis. 17, 1970-71.
619 - Kahankova, J., Pantucek, R., Goerke, C., Ruzickova, V., Holochova, P. et al., 2010.
620 Multilocus PCR typing strategy for differentiation of *Staphylococcus aureus* siphoviruses
621 reflecting their modular genome structure. Environ. Microbiol. 12, 2527-38.
622 - Kaneko, J., Kimura T., Narita S., Tomita, T., Kamio, Y., 1998. Complete nucleotide
623 sequence and molecular characterization of the temperate staphylococcal bacteriophage
624 ϕSLT carrying Panton-Valentine leukocidin genes. Gene. 215, 57-67.
625 - Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H. et al., 2001. Whole genome
626 sequencing of meticillin-resistant *Staphylococcus aureus.* Lancet. 357, 1225-40.
627 - Lee JC., 1995. Electrotransformation of staphylococci. Methods Mol. Biol. 47, 209-216.
628 - Marraffini, L.A., and Sontheimer, E.J., 2008. CRISPR interference limits horizontal gene
629 transfer in staphylococci by targeting DNA. Science. 322, 1843-45.

630 - McCarthy, A.J., Witney, A.A., Gould, K.A., Moodley, A., Guardabassi L. et al., 2011. The
631 distribution of MGEs in MRSA CC398 is associated with both host and country. Genome
632 Biol. Evol. 3, 1164-74.
633 - Messad, N., Lina, G., Prasjnar, T.K., Renshaw, S.A., Vandenesch, F. et al. 2012. Insertion
634 d'un phage dérégulant la virulence de *Staphylococcus aureus.* Abstr. 32ième Réunion
635 Interdisciplinaire de Chimiothérapie Anti-Infectieuse. O-15.
636 - Nair, D., Memmi, G., Hernandez, D., Bard, J., Beaume, M. et al., 2011. Whole-genome
637 sequencing of *Staphylococcus aureus* strain RN4220, a key laboratory strain used in
638 virulence research, identifies mutations that affect not only virulence factors but also the
639 fitness of the strain. J. Bacteriol. 193, 2332-35.
640 - Park, P.W., Senior, R.M., Griffin, G.L., Broekelmann, T.J., Mudd, M.S., Mecham, R.P.,
641 1999. Binding and degradation of elastin by the staphylolytic enzyme lysostaphin. Int. J.
642 Biochem. Cell. Biol. 27, 139–146.
643 - Price, L.B., Stegger, M., Hasman, H., Aziz, M., Larsen, J. et al., 2012. *Staphylococcus*
644 *aureus* CC398: Host Adaptation and Emergence of Methicillin Resistance in Livestock.
645 MBio. 3, e00305-11.
646 - Rashel, M., Uchiyama, J., Ujihara, T., Uehara, Y., Kuramoto, S. et al., 2007. Efficient
647 elimination of multidrug-resistant *Staphylococcus aureus* by cloned lysin derived from
648 bacteriophage ϕMR11. J. Infect. Dis. 196, 1237-47.
649 - Rooijakkers, S.H., Ruyken, M., van Roon, J., van Kessel, K.P., van Strijp, J.A. el al., 2006.
650 Early expression of SCIN and CHIPS drives instant immune evasion by *Staphylococcus*
651 *aureus*. Cell. Microbiol. 8, 1282-93.

- Schijffelen, M.J., Boel, C.H., van Strijp, JA., 2010. Whole genome analysis of a livestock-associated methicillin-resistant *Staphylococcus aureus* ST398 isolate from a case of humen endocarditis. BMC. Genom. 11, 376.
- Sinha, B., François, P.P., Nüsse, O., Foti, M., Hartford, O.M. et al., 1999. Fibronectin-binding protein acts as *Staphylococcus aureus* invasin via fibronectin bridging to integrin alpha5beta1. Cell. Microbiol. 1, 101-17.
- Smeltzer, M.S., Hart, M.E., Iandolo J.J., 1994. The effect of lysogeny on the genomic organization of *Staphylococcus aureus*. Gene. 138, 51-7.
- Stegger, M., Lindsay, J.A., Sorum, M., Gould, K.A., Skov, R., 2010. Genetic diversity in CC398 methicillin-resistant *Staphylococcus aureus* isolates of different geographical origin. Clin. Microbiol. Infect. 16,1017-19.
- Stegger, M., Lindsay, J.A., Moodley, A., Skov, R., Broens, E.M. et al. 2011. Rapid PCR detection of *Staphylococcus aureus* clonal complex 398 by targeting the restriction-modification system carrying sau1-hsdS1. J Clin. Microbiol. 49,732-4.
- Sumby, P. and Waldor, M.K., 2003. Transcription of the toxin genes present within the staphylococcal phage phiSa3ms is intimately linked with the phage's life cycle. J. Bacteriol. 185, 6841-51.
- Ubeda, C., Maiques, E., Knecht, E., Lasa, I., Novick, R.P., Penades, J.R., 2005. Antibiotic-induced SOS response promotes horizontal dissemination of pathogenicity island-encoded virulence factors in staphylococci. Mol. Microbiol. 56, 836-44.
- Uhleman, A.C., Porcella, S., Trivedi, S., Sullivan, S., Hafer, C. et al., 2012. Identification of a highly transmissible animal-independent *Staphylococcus aureus* ST398 clone with distinct genomic and cell adhesion properties. Mbio. 3, e00027-12.

- Valentin-Domelier, A.S., Girard, M., Bertrand, X., Violette, J., François, P. et al., 2011. Methicillin-susceptible ST398 *Staphylococcus aureus* responsible for bloodstream infections: an emerging human-adapted subclone ? PloS One. 6, e28369.
- van Belkum, A., Melles, D.C., Peeters, J.K., van Leeuwen, W.B., van Duijkeren, E. et al., 2008. Methicillin-resistant and -susceptible *Staphylococcus aureus* sequence type 398 in pigs and humans. Emerg. Infect. Dis. 14, 479-83.
- van Wamel, W.J., Rooijakkers, S.H., Ruyken, M., van Kessel, K.P., van Strijp, J.A., 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, 1310-15.
- Voyich, J.M., Braughton, K.R., Sturdevant, D.E., Whitney, A.R., Porcella, S.F. et al., 2005. Insights into mechanisms used by *Staphylococcus aureus* to avoid destruction by human neutrophils. J. Immunol. 175, 3907-19.
- Wagner, P.L., Waldor, M.K., 2002. Bacteriophage control of bacterial virulence. Infect. Immun. 70, 3985-93.
- Waldron, D.E., Lindsay, J.A., 2006. Sau1: a novel lineage-specific type I restriction-modification system that blocks horizontal gene transfer into *Staphylococcus aureus* and between *S. aureus* isolates of different lineages. J. Bacteriol. 188, 5578-85.
- Witte, W., Strommenger, B., Stanek, C., Cuny, C., 2007. Methicillin-resistant *Staphylococcus aureus* ST398 in humans and animals, Central Europe. Emerg. Infect. Dis. 13, 255-58.
- Xia, G., Wolz, C. 2013. Phages of *Staphylococcus aureus* and their impact on host evolution. Infect. Genet. Evol. pii: S1567-1348(13)00171-8. doi: 10.1016/j.meegid.2013.04.022.

- Yamaguchi, T., Hayashi, T., Takami, H., Nakasone, K., Ohnishi, M. et al., 2000. Phage conversion of exfoliative toxin A production in *Staphylococcus aureus*. Mol. Microbiol. 38, 694-705.

FIGURE LEGENDS

Figure 1.

Microarray analysis of LA and non-LA CC398 isolates. Clusters of hybridization probes that distinguish LA from non-LA *S. aureus* CC398. Black dots show probes displaying positive fluorescent signals.

Figure 2.

Schematic representation of the recombination between StauST398-4pro and StauST398-5pro that led to StauST398-1. Note that most of the sequence originates from StauST398-5 and is flanked by StauST398-4 extremities. This recombination resulted in the partial deletion of the IEC locus.

Figure 3.

Epithelial cell invasion experiments**.** Invasiveness of *S. aureus* strains in 293 cells was determined relative to that of the positive control strain (Cowan). Total internalized bacteria were enumerated by counting colony forming units following lysostaphin treatment. The values are means ± SEM of four independent experiments.

Figure 4.

Expression of phage-encoded virulence genes following various stresses**.** Three lysogenic strains representative of the non-LA CC398 isolates [S94, subgroup A, harboring Stau398-4pro; S81, subgroup E, harboring Stau398-4pro and Stau398-5pro, and the MRSA S89] were subjected to stresses during exponential growth or the stationary phase, including sublethal concentrations of oxacillin (OXA), erythromycin (ERY), tetracycline (TETRA) or pestanal.

726 Levels of *chp* mRNA (left panel) and *scn* mRNA (right panel) were assessed in strains S94
727 (white bars), S89 (grey bars) and S81 (black bars) by RT-qPCR. Results are mean values for
728 two independent measurements performed in triplicate (mean ± range).

Table 1.

Isolation context, antibiotic resistance, *spa*-type, phage susceptibility, phage profiling and IEC-typing of the 21 CC398 isolates and transductant S124/S100 ϕ.

	Characteristics of CC398 isolates																						
	LA[1]										non-LA[1]												
											subgroup A							subgroup E					
	S123	S124	S132	S131	S1	S129	S125	S126	S134	S130	S92	S94	S93	S106	S103	S91	S89	S105	S96	S81	S100	S124 /S100-ϕ[2]
Hcol / Hinf / A[3]	A	A	A	A	Hcol	Hcol	A	Hcol	A	Hcol	Hinf	Hinf	Hinf	Hinf	Hinf	Hinf	Hinf	Hinf	Hinf	Hinf	Hinf	
Spa-type	11	34	34	11	571	108	108	571	899	899	5635	571	1451	571	571	571	899	1451	6605	571	9378	34
Antibiotic resistance[4]	Tet Ery	Tet	oxa Tet KT Ery	oxa Tet	oxa Tet GnT Ery	oxa Tet	Tet Ery	oxa Tet	oxa Tet	oxa Tet KT Ery	Ery	Ery	Ery	Ery	Ery		oxa Tet		Ery	Ery		Tet
Susceptibility to phages	++	++	no	no	no	no	no	no	+	no	no	no	no	no	no	no	no	no	no	no	no	no
Prophage profiling[5]				int2	int2	int2	int2		int2	int2												
lysogeny			int6					int6		int6	int3	int3	int3	int3	int3	int3	int3	int3	int3	int3	int3	int3
			ant1b		ant1b					ant1b	ant4b	ant4b	ant4b	ant4b	ant4b	ant4b	ant1b	ant4b	ant4b	ant4b	ant4b	ant4b
				ant5				ant5														
replication										dnaC1	dnaD1b	dnaD1b	dnaD1b	dnaD1b	dnaD1b	dnaD1b		dnaC1 dnaD1b	dnaC1 dnaD1b	dnaC1 dnaD1b	dnaC1 dnaD1b	dnaC1 dnaD1b
			dnaD2b	dnaD2b	dnaD2b												dnaD2b					
					polA	polA	polA	polA	polA	polA												
regulation of transcription			dut2	dut2				dut2		dut2								dut2	dut2	dut2	dut2	dut2
					dut3												dut3					
morphogenesis subtype											Fa	Fa	Fa	Fa	Fa	Fa	Fa	Fa	Fa	Fa	Fa	Fa
					Bb					Bb								Bb	Bb	Bb	Bb	Bb
																	Be					
phage type		A	A	A	A	A	A	A	A	A												
					B					B												
	F	F					F			F	F	F	F	F	F	F	F	F	F	F	F	F
lytic module					ami1			ami1									ami1					
			ami2	ami2		ami2	ami2		ami2	ami2	ami2	ami2	ami2	ami2	ami2	ami2		ami2	ami2	ami2	ami2	ami2
										ami3							ami3					
																		ami4	ami4	ami4	ami4	ami4
IEC (genes present)	no	no	no	no	no	no	no	no	no	no	+ (*chp scn*)	+ (*chp scn*)	+ (*chp scn*)	+ (*chp scn*)	+ (*chp scn*)	+ (*chp scn*)	+ (*chp scn*)	+ (*chp scn*)	+ (*chp scn*)	+ (*chp scn sak*)	+ (*chp scn*)	+ (*chp scn*)

[1]LA: livestock-associated, non-LA: from an animal-free environment, [2]transductant obtained with infective lysate S100-ϕ used to transduce the prophage-free isolate S124. [3]Hcol: human colonizing isolate, Hinf: clinical isolate recovered from an invasive infection, A: animal colonizing isolate, [4]oxa: oxacillin, Tet: tetracyclin, Ery: erythromycin, K: kanamycin, T: tobramycin, [5]The

genomic modules were classified on the basis of the genes for integrase (10 types), anti-repressor (five types), replication proteins *polA*, *dnaC* and *dnaD* (four types), dUTPase (four types), portal protein (eight types), tail appendices (four types) and endolysin (four types) corresponding to the integrase locus, lysogeny control region, and modules for DNA replication, transcription regulation, packaging, tail appendices and lysis respectively.

743 **Table 2.**

744 Transformation efficiencies (number of transformants/µg of plasmid DNA)

745

Recipient isolate				Transformation efficiency according to transforming DNA (number of transformants/µg of plasmid DNA)[1]		
	Origin	Prophage content	GenBank Accession numbers	*E. coli*DH5α	RN4220	*E. faecalis*[3]
RN4220				107050	172543	2-11
S123	LA	No		4175	1011	1-4
S124	LA	no		5048	818	ND
S100	human	StauST398-4pro + StauST398-5pro	KC595278 and KC595279	2	<1	0
S124/S100-ϕ [2]		StauST398-1pro	JX013863	401	41	ND
S92	human	StauST398-4pro		90	14	2
S93	human	StauST398-4pro		90	24	ND
S94	human	StauST398-4pro		463	30	1

746

747 [1] Values (mean of 2 experiments) represent the number of transformants per µg of plasmid DNA, normalized to

748 10^{10}electroporated cells.

749 [2] Results affected by the instability of StauST398-1

750 [3] Values obtained for 2 experiments.

751

752

753

754