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

To cite this version:
Nathalie Van Der Mee-Marquet, Anna-Rita Corvaglia, Anne-Sophie Valentin, David Hernandez, Xavier Bertrand, et al.. Analysis of prophages harbored by the human-adapted subpopulation of Staphylococcus aureus CC398. Infection, Genetics and Evolution, Elsevier, 2013, 18, pp.299-308. <10.1016/j.meegid.2013.06.009>. <hal-00843111>
Analysis of prophages harbored by the human-adapted subpopulation of Staphylococcus aureus CC398

Nathalie van der Mee-Marquet¹,²*, Anna-Rita Corvaglia³, Anne-Sophie Valentin¹, David Hernandez³, Xavier Bertrand⁴,⁵, Myriam Girard³, Jan Kluytmans⁶, Pierre-Yves Donnio⁷,⁸, Roland Quentin¹, Patrice François³

¹Service de Bactériologie et Hygiène, Centre Hospitalier Régional Universitaire, and ²Réseau des Hygiénistes du Centre, Centre Hospitalier Universitaire, F37044 Tours cedex, France. ³Genomic Research Laboratory, University of Geneva Hospitals, CH-1211 Geneva 14, Switzerland. ⁴Service d'Hygiène, Centre Hospitalier Universitaire, and ⁵UMR 6249 Chrono-environnement, Université de Franche-Comté, F25030 Besançon cedex, France. ⁶Laboratory for Microbiology and infection Control, Amphia Hospital, Breda and VU University Medical Center, Amsterdam, The Netherlands. ⁷Service de Bactériologie et Hygiène Hospitalière, Centre Hospitalier Universitaire, and ⁸EA1254 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.
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 *hbl*, *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
previously described (Beaume, 2010). The media used for overnight culture and subsequent mini-preps (Qiagen) included Luria Bertani Broth (LB Broth, Miller, BD Difco) for *E. coli* DH5α, supplemented with 100 mg/L of ampicillin; Mueller-Hinton broth (MHB; BD Difco) for *S. aureus* RN4220; and Tryptic soy Broth for *E. faecalis* (Soybean-Casein Digest Medium, BD Bacto). Chloramphenicol (10 mg/L) was included as appropriate. Preparation of competent cells and transformation of RN4220, S100, S123, S124, S124/S100-φ, S92, S93 and S94 were as described previously (Beaume, 2010). Following electroporation, transformants were plated on Mueller-Hinton Agar, supplemented with chloramphenicol: 10 mg/L for RN4220; 15 mg/L for S100, S123, S124, S124/S100-φ; and 20 mg/L for S92, S93, S94. The number of transformants obtained per µg of plasmid DNA was calculated. **Internalization experiments.** The Cowan strain is internalized by human cells and was therefore used as a positive control. Strain KH11 was used as a non-invasive, negative control. The internalization procedure and the lysostaphin protection assay were based on a published method (Sinha, 1999) with minor differences. Briefly, cultures of the cell line 293 were maintained under humidified air with 5 % CO₂ at 37°C in Dulbecco’s minimal essential medium (DMEM)-nutrient mixture F-12 (nut mix F-12) containing Glutamax I, a stable glutamine dipeptide, supplemented with 10 % fetal calf serum (FCS), 50 IU of penicillin per mL, and 50 µg of streptomycin per mL. For the lysostaphin assay, 293 cells were plated in poly-ornithine coated plates (1.8 cm diameter) and cultured for 24 hours, such that there were approximately 0.9 x 10⁵ cells/well. Bacterial were harvested from culture, washed, counted and diluted (Sinha, 1999); these bacterial dilutions were added to the plates containing 293 cells at a multiplicity of infection [MOI] of 15:1. The plates were incubated for 1 h at 37°C, and the bacterial suspensions were then replaced with lysostaphin medium (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$_2$O. 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.
RESULTS

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

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

Prophage profiling and microarray data revealed two subgroups among the non-LA MSSA isolates (Tab.1; Fig.1). One subgroup comprised six isolates only harboring Sa3 prophage elements. Four isolates in the second subgroup harbored additional prophage elements next to the Sa3 prophage elements. These additional elements hybridized with 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 and 138nm-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
known restriction system was found. However, part of a putative type III restriction-modification system was identified by the annotating tool. The phage genome carries two putative virulence genes, *chp* and *scn*, around 2.5 Kb upstream from the insertion site. The sequence of StauST398-4pro shares approximately 97% identity with phi3, recently described in *S. aureus* strain 71193 (Uhleman, 2012) and 91% with the close neighbor of phiNM3, recently described by Price and colleagues (Price, 2012). The two other phages described by Price and colleagues are much more divergent from StauST398-4pro (<15% identity). PCR with primers specific for StauST398-4pro, amplified fragments of this β-converting φ3-prophage from all of the non-LA CC398 isolates but none of the LA isolates (Suppl. File Tab.2).

The StauST398-5pro genome (accession number KC595279) is a 40.2 Kb, linear, double-stranded DNA with 0.353% G+C. It carries 55 putative coding sequences, including a putative gene for a superantigen similar to enterotoxin B. It is inserted in *smpB*. It has a modular organization with gene clusters involved in DNA packaging, head and tail morphogenesis, cell lysis, lysogeny, replication and regulation. The lysogeny module was found between the lysis cassette and the replication module, with a repressor protein cI, the repressor cro and an antirepressor that appeared to result from antA/B recombination (Davis, 2002). The transcriptional regulation module comprises a dimeric dUTPase, an integrase activator and a rho terminator factor. This phage lacks an integrase, indicating a lytic-defective prophage. Homology searches identified a putative restriction system. StauST398-5pro shares 75% identity with φMR11 (Rashel, 2007) and with φETA (Yamaguchi, 2000) over more than half of its length. The remaining 15 Kb of the genome showed a highly 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 $\beta$-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,
but also a φMR11-like prophage (named StauST398-5pro). StauST398-4pro is an IEC β-
converting Sa3 prophage that has been described in isolates responsible for human
infections in LA- and non-LA environments (Stegger, 2010; Valentin-Domelier, 2011;
Jimenez, 2011), but StauST398-5pro has not previously been detected in isolates responsible
for human infections in a LA environment. Therefore, we suggest that isolates only harboring
StauST398-4pro belong the ancestral human LA subpopulation, and that StauST398-5pro is
specific to isolates responsible for the more recent human invasive infections in a non-LA
environment, i.e. isolates of the emerging human-adapted non-LA CC398 subpopulation.

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

We demonstrate that CC398 prophages, like most described S. aureus prophages,
enabled their hosts to exclude other phages and exogenous DNA. However, the protection
conferred by these prophages varied: the protection conferred by StauST398-4pro prophage
was however limited, but if StauST398-5pro prophage was present in the non-LA isolate, the
protection was strong. *In silico* analysis of complete genomes from naturally lysogenic non-LA CC398 did not identify any CRISPR or CRISPR-associated genes. Prophage sequences encoding a putative RM system were identified in StauST398-5pro. Thus, the emerging CC398 isolates, all of which harbor StauST398-5pro, have a prophage-encoded RM system conferring strong protection against horizontal genetic transfer, similar to that described for the φ42 prophage (Dempsey, 2005). This protection may confer a selective advantage over competing organisms in their environment.

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

Prophages can contribute directly to bacterial pathogenesis at the time of infection (Sumby, 2003). The expression of temperate prophage genes, that are generally repressed,
may be induced during stress and/or induction of the SOS system, because the repressor, cl, is inactivated by RecA. This can result in increased transcription of virulence factors, as has been demonstrated for staphylokinase production in *S. aureus* strains carrying β-converting phages with the *sak* gene (Rooijakkers, 2006). We demonstrated that the non-LA prophage StauST398-5pro encodes a putative superantigen similar to enterotoxin B and the non-LA β-converting prophage StauST398-4pro encodes and expresses two human-specific virulence genes (*chp* and *scn*) (van Wamel, 2006). *S. aureus* enters the body through breaches in the skin or mucous membranes, and is then immediately confronted by the host innate defense system. To counteract innate immunity, *S. aureus* expresses immune-modulating proteins encoded by β-converting prophages, that facilitate long-term colonization in humans (Goerke, 2006; Rooijakkers, 2006): the staphylococcal complement inhibitor (SCIN) prevents opsonophagocytosis and killing of *S. aureus* by human neutrophils, the chemotaxis inhibitory protein of *S. aureus* (CHIPS), the bacterial plasminogen activator (SAK) and the superantigens SEA and SEB, that down-regulate chemokine receptors on monocytes. Our transcription experiments indicate that non-LA CC398 lysogens may benefit from the production of SEB, CHIPS and SCIN (products of *seb*, *chp* and *scn* genes, respectively) when exposed to conditions that favor prophage induction.

In the genome of StauST398-5pro, we identified a transcription repressor and antirepressor resulting from antA/B recombination, characteristic of a helper prophage (Davis, 2002). The association between a helper-phage and a satellite prophage allows the production proteins encoded by the satellite phage that may escape lysogenic repression because of interactions between repressors and antirepressors. In *S. aureus*, the association between a β-converting phage and a helper phage has been shown to result in the 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 StauST398-5-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 lose 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...
content can influence the phenotypes that contribute to genome plasticity (notably transformation and transduction) and to the invasiveness of *S. aureus*. These prophages in the bacterial host provide protection against horizontal genetic transfer, and also contribute to virulence and adaptation of the host bacterium.
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), FOUJNIER-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), HOMBREUOCK-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), NARBAY 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


- Nair, D., Memmi, G., Hernandez, D., Bard, J., Beaume, M. et al., 2011. Whole-genome sequencing of *Staphylococcus aureus* strain RN4220, a key laboratory strain used in virulence research, identifies mutations that affect not only virulence factors but also the fitness of the strain. J. Bacteriol. 193, 2332-35.


- Sumby, P. and Waldor, M.K., 2003. Transcription of the toxin genes present within the staphylococcal phage philSa3ms is intimately linked with the phage’s life cycle. J. Bacteriol. 185, 6841-51.


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.
Levels of *chp* mRNA (left panel) and *scn* mRNA (right panel) were assessed in strains S94 (white bars), S89 (grey bars) and S81 (black bars) by RT-qPCR. Results are mean values for 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.

<table>
<thead>
<tr>
<th>Characteristics of CC398 isolates</th>
<th>subgroup A</th>
<th>non-LA</th>
<th>subgroup E</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA 1</td>
<td>S123</td>
<td>S124</td>
<td>S132</td>
</tr>
<tr>
<td>Spa-type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antibiotic resistance 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Susceptibility to phages</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prophage profiling / lysogeny</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>replication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>regulation of transcription</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>morphogenesis subtype</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>phage type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lytic module</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IEC (genes present)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

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

<table>
<thead>
<tr>
<th>Recipient isolate</th>
<th>Transformation efficiency according to transforming DNA (number of transformants/μg of plasmid DNA)¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Origin Prophage content</td>
<td>GenBank Accession numbers</td>
</tr>
<tr>
<td>RN4220 S123</td>
<td>LA</td>
</tr>
<tr>
<td>RN4220 S124</td>
<td>LA</td>
</tr>
<tr>
<td>RN4220 S100</td>
<td>human</td>
</tr>
<tr>
<td>+ StauST398-5pro and KC595279</td>
<td></td>
</tr>
<tr>
<td>RN4220 S124/S100-φ²</td>
<td>StauST398-1pro</td>
</tr>
<tr>
<td>RN4220 S92</td>
<td>human</td>
</tr>
<tr>
<td>RN4220 S93</td>
<td>human</td>
</tr>
<tr>
<td>RN4220 S94</td>
<td>human</td>
</tr>
</tbody>
</table>

¹ Values (mean of 2 experiments) represent the number of transformants per μg of plasmid DNA, normalized to 10¹⁰ electroporated cells.

² Results affected by the instability of StauST398-1

³ Values obtained for 2 experiments.