Diversity of Prophages in Dominant *Staphylococcus aureus* Clonal Lineages[▽]

Christiane Goerke,^{1*} Roman Pantucek,² Silva Holtfreter,³ Berit Schulte,¹ Manuel Zink,¹
Dorothee Grumann,³ Barbara M. Bröker,³ Jiri Doskar,² and Christiane Wolz¹

Institut für Med. Mikrobiologie und Hygiene, Universitätsklinikum Tübingen, Tübingen, Germany¹; Department of Genetics and Molecular Biology, Masaryk University, Brno, Czech Republic²; and Institut für Immunologie und Transfusionsmedizin, Universität Greifswald, Greifswald, Germany³

Received 23 December 2008/Accepted 23 March 2009

Temperate bacteriophages play an important role in the pathogenicity of *Staphylococcus aureus*, for instance, by mediating the horizontal gene transfer of virulence factors. Here we established a classification scheme for staphylococcal prophages of the major *Siphoviridae* family based on integrase gene polymorphism. Seventy-one published genome sequences of staphylococcal phages were clustered into distinct integrase groups which were related to the chromosomal integration site and to the encoded virulence gene content. Analysis of three marker modules (lysogeny, tail, and lysis) for phage functional units revealed that these phages exhibit different degrees of genome mosaicism. The prevalence of prophages in a representative *S. aureus* strain collection consisting of 386 isolates of diverse origin was determined. By linking the phage content to dominant *S. aureus* clonal complexes we could show that the distribution of bacteriophages varied remarkably between lineages, indicating restriction-based barriers. A comparison of colonizing and invasive *S. aureus* strain populations revealed that *hly*-converting phages were significantly more frequent in colonizing strains.

Staphylococcus aureus asymptotically colonizes the anterior nares of humans but also causes a wide spectrum of acute and chronic diseases. Most of the dissimilarity between *S. aureus* strains is due to the presence of mobile genetic elements such as plasmids, bacteriophages, pathogenicity islands, transposons, and insertion sequences (2, 14, 19, 23). Many virulence factors are encoded on such mobile elements (3, 6, 17, 26, 27, 35). In particular, bacteriophages play an important role in the pathogenicity of *S. aureus* either by carrying accessory virulence factors such as Panton-Valentine leukocidin (PVL) (encoded by the *luk-PV* operon), staphylokinase (encoded by *sak*), enterotoxin A (encoded by *sea*), and exfoliative toxin A (encoded by *eta*) or by interrupting chromosomal virulence genes such as those for β -hemolysin (*hly*) and lipase (*geh*) upon insertion. Additionally, phages are the primary vehicle of lateral gene transfer between *S. aureus* strains, providing the species with the potential for broad genetic variation. We could show that phages increase the genome plasticity of *S. aureus* during infection, facilitating the adaptation of the pathogen to various host conditions (11, 12).

Despite the obvious importance of phages for the biology of *S. aureus*, epidemiological data on the prevalence of phages in this species are limited (28, 33). More than 80 genome sequences of staphylococcal bacteriophages and prophages are available in the public genome databases. Most published *S. aureus* phages belong to the *Siphoviridae* family of temperate, tailed bacterial viruses. Traditionally, *S. aureus* phages were characterized according to their lytic activity, morphology, and

serological properties (1, 28). Today, the temperate phages in clinical *S. aureus* isolates can be identified with a multiplex PCR strategy, which is based on sequence differences between viral genes coding for the surface-exposed determinants (28).

In general, the evolution of phage lineages seems to be driven by the lateral gene transfer of interchangeable genetic elements (modules), which consist of functionally related genes. The *Siphoviridae* genomes are usually organized into six functional modules: lysogeny, DNA replication, regulation of transcription, packaging and head, tail, and lysis (4). A functional module found in one phage can be replaced in another phage by a sequence-unrelated module that fulfills the same or related functions. Multiple alignment of *S. aureus* phage genomes also revealed a chimeric and mosaic structure resulting from horizontal transfer and recombination (5, 20). It is an open question whether all phages have access to a common gene pool or whether subpools have developed, which are due to differences in the accessibility of strain variants of the bacterial host species.

It was recently shown that most human *S. aureus* strains belong to one of 10 independent lineages or clonal complexes (CCs) (9, 24). Exchange of DNA is very much lower between different lineages than within the same lineage due to the action of the restriction-modification (R-M) system *SauI* (34). The prime role of R-M systems in many bacteria is the defense against DNA bacteriophages. These systems usually comprise a DNA methyltransferase and a restriction endonuclease. The former protects self DNA by methylation of specific nucleotides in a certain DNA sequence, whereas the latter cleaves the foreign unmodified DNA at the same sequence motif. If the spread of bacteriophages between different *S. aureus* lineages was controlled by the R-M system or a similar mechanism, an unequal distribution would be expected.

Here we established a classification scheme for staphylococ-

* Corresponding author. Mailing address: Institut für Med. Mikrobiologie und Hygiene, Universitätsklinikum Tübingen, Elfriede-Aulhorn-Str. 6, 72076 Tübingen, Germany. Phone: 49-7071-2985229. Fax: 49-7071-295165. E-mail: christiane.goerke@med.uni-tuebingen.de.

[▽] Published ahead of print on 27 March 2009.

cal prophages of the major *Siphoviridae* family which was based on the suggested phage designation of the published *S. aureus* genomes (22, 23). When analyzing a representative *S. aureus* strain collection, we could show that the frequency of certain phage groups varied between *S. aureus* lineages. A comparison of colonizing and invasive *S. aureus* strain populations revealed that *hly*-converting phages were significantly more frequent in colonizing strains.

MATERIALS AND METHODS

Bacterial isolates. *S. aureus* isolates from different sources were included in this study: nasal carriage isolates and blood culture isolates from the University of Greifswald (15); nasal carriage isolates (12) and clinical isolates from diverse sources from the Department of Medical Microbiology and Hygiene, Tübingen; blood culture isolates from the Kantonsspital Basel; community-associated methicillin-resistant *S. aureus* (MRSA) reference isolates from the Robert Koch Institute, Wernigerode; and reference strains from the Network on Antimicrobial Resistance in *Staphylococcus aureus* strain collection.

Strain typing. *spa* typing was performed as described previously (15) using the Ridom StaphType software (13). Pulsed-field gel electrophoresis (PFGE) typing and Southern blot analysis were performed as described previously (10).

MLST. Multilocus sequence genotyping (MLST) was performed on selected isolates according to published protocols (8). Otherwise, MLST CCs were deduced from BURP grouping of *spa* types or by comparison with the PFGE pattern (31).

Phage integrase multiplex PCRs. Multiplex PCR was performed using the multiplex PCR kit (Qiagen, Hilden, Germany). Each reaction mixture (25 μ l) contained 12.5 μ l 2 $\times$ Qiagen multiplex PCR master mix, 200 nM of each primer, and 10 ng of template DNA. An initial denaturation of DNA at 95°C for 15 min was followed by 35 cycles of amplification (95°C for 30 s, 55°C for 60 s, and 72°C for 45 s), ending with a final extension phase at 72°C for 10 min. All PCR products were resolved by electrophoresis in 3% agarose gels, stained with ethidium bromide, and visualized under UV light. The following primers specific for the phage integrase genes were used: for Sa1int, Sa1-F (AAGCTAAGTTCGGGCACA) and Sa1-R (GTAATGTTTGGGAGCCAT) (length, 569 bp); for Sa2int, Sa2-F (TCAAGTAACCCGTCACCTC) and Sa2-R (ATGTCTAAATGTGTGCGTG) (length, 640 bp); for Sa3int, Sa3-F (GAAAAACAAACGGGTGCTAT) and Sa3-R (TTATTGACTCTACAGGCTGA) (length, 475 bp); for Sa4int, Sa4-F (ATTGATATTAACGGAACCTC) and Sa4-R (TAAACTTATATGCTGTGT) (length, 320 bp); for Sa5int, Sa5-F (AAAGATGCCAACTAGCTG) and Sa5-R (CTTGTGGTTTTGTTCTGG) (length, 375 bp); for Sa6int, Sa6-F (GCCATCAATTCAAGGATAG) and Sa6-R (TCTGCAGCTGAGGACAAT) (length, 167 bp); and for Sa7int, Sa7-F (GTCCGGTAGCTAGAGGTC) and Sa7-R (GGCGTATGCTTGACTGTGT) (length, 214 bp). Validation of the multiplex PCR assay was carried out with (i) genome-sequenced *S. aureus* strains, (ii) phage-less *S. aureus* 8325-4 or *S. aureus* 1039 lysogenized with genome-sequenced phages of the International Typing Set belonging to different *int* gene classes (ϕ 55, Sa1int; ϕ 47, Sa2int; ϕ 42E, Sa3int; ϕ 29, Sa5int; ϕ 77, Sa6int; or ϕ 53, Sa7int), and (iii) triple-lysogenic *S. aureus* NCTC 8325 (harboring prophages ϕ 11, ϕ 12, and ϕ 13) lysogenized with ϕ 77 or with ϕ 53.

Sequence analysis. Phage sequences were obtained from the NCBI nucleotide database or were assembled from the published *S. aureus* genomes. Open reading frames (ORFs) for the integrase and holin genes were deduced from the whole phage genomes by BLAST analysis. Integrase sequences were aligned by ClustalW. Phages were assigned to serotypes using the primer sequences published by Pantucek et al. to identify the respective genes (28).

For sequencing of the integrase and holin genes of phage ϕ 6390, the targets were amplified from strain RN6390 by standard PCR using the primers 6390intseq-for (ATTGGCGAACGAGGTAAC) and 6390intseq-rev (GCCAATTTGAGGAGGGAG) for the integrase gene and Holin255-for (ATGATGAATTGGAAAATTAGAA) and Holin255-rev (CTAGTATTTTCTTCTTGGTTCT) for the holin gene. Amplicons were cloned into pCR2.1 (Invitrogen, Karlsruhe, Germany) for sequencing. The sequencing was done by 4base lab, Reutlingen, Germany, using the Dynamic sequence kit (Amersham Biosciences, Freiburg, Germany). Sequence data were analyzed using Vector NTI software (InforMax, Frederick, MD).

Statistical analysis. Differences between groups were assessed using the likelihood ratio test with Bonferroni's adjustment.

Nucleotide sequence accession number. The ϕ 6390 integrase sequence was deposited in the GenBank and EMBL databases (accession no. FM877489).

RESULTS

Classification of staphylococcal prophage modules. The identification and description of prophages existing in bacterial strains relies on a clear classification scheme, while the grouping of bacteriophages into distinct phage types is extremely difficult because of high sequence variation even within functionally equivalent genes. Additionally, recombination leads to the emergence of extensive mosaicism in bacteriophage genomes (5, 20). To develop a reliable classification system, we compared all 71 complete bacteriophage genome sequences of the class *Siphoviridae* currently available in the databases; the majority are derived from *S. aureus*, and two each are from *Staphylococcus epidermidis* and *Staphylococcus haemolyticus*. The comparison was based on three distinct features present in all phages: (i) differences in the lysogeny module, in particular the integrase, which dictates the insertion site of the phage in the bacterial chromosome; (ii) differences in the phage morphology module, where the serogroup of each phage was determined based on capsid, tail, and tail appendix protein sequences (28), and (iii) differences in the lysis module, where the characteristic holin gene was investigated.

Genes coding for a putative integrase could be identified in all the available staphylococcal phage genomes with the exception of one (ϕ 3A). Alignment of the integrase nucleotide sequences clustered the prophages in seven major and eight minor groups (Fig. 1). Within groups, the nucleotide sequence identity was 95% to 100%; between the groups, there was only 38% to 84% identity, which was still compatible with functional analogy. In fact, based on amino acid sequence homology and catalytic residues, most integrases belong to the tyrosine recombinase type family; only Sa7int, Sa12int, Se1int, and Sh2int belong to the serine recombinase type family. Most *S. aureus* prophages clustered in one of the seven major groups (designated Sa1int to Sa7int), and five were singletons (Sa8int to Sa12int). The two known *S. epidermidis* bacteriophages (ϕ CNPH82 and ϕ PH15) and the two *S. haemolyticus* prophages (ϕ JCSC1435A and ϕ JCSC1435B) differed strongly from all *S. aureus* phages (Fig. 1).

Next, the serogroup-specifying genes were compared based on capsid, tail, and tail appendix protein sequences (28). The three main serogroups A, B, and F were associated with phage tail appendices. F phages were classified into two subgroups, Fa and Fb, because their DNA-packaging, head, and tail genes belong to different modules. Sequence homology within groups ranged from 85% to 100%, while there was no significant homology between the groups. The majority of *S. aureus* phages could be assigned to one of the four prominent serogroups (A, B, Fa, and Fb). Two *S. aureus* phages (ϕ 37 and ϕ EW), the two *S. epidermidis* phages, and the two *S. haemolyticus* phages could not be discerned with the applied classification scheme.

Sequence alignment of the holin genes revealed 10 different groups. Sequence relatedness within these groups ranged from 92% to 100% homology. Because sequence homology was closely correlated with gene length, the holin groups were designated by the sequence length polymorphism (number), and additional sequence variations were indicated by letters (255a, 255b, 216, 273, 276, 303, 423, 435, 438, and 486). The majority of the *S. aureus* phages clustered in one of the five

major holin groups. Little sequence homology was observed between the holin genes from the different staphylococcal species.

Mosaicism of staphylococcal phage genomes. When comparing the integrase sequence tree with the results of the other two classification systems for 70 staphylococcal phages (ϕ 3A was omitted, because no integrase gene could be detected), different degrees of genome mosaicisms were observed. For the phage groups Sa3int, Sa5int, and Sa6int, a high diversity in the combination of the three marker modules was determined (Fig. 1). The largest group of *S. aureus* phages (14/70) are the Sa3int phages, which differed in serogroup (Fa, Fb, or A), holin genes (255a and 255b are not closely related), and their combination of immune-modulatory virulence factors. Phages of this group typically integrate into the *hly* gene of *S. aureus*, leading to negative conversion of β -hemolysin production (6). Similarly, the phages of the Sa5int group exhibited a high diversity in their module pattern: three serogroups (B, Fb, and L) and three holin groups (255a, 303, and 438) could be detected. The popular transducing phage ϕ 11 of *S. aureus* strain 8325 is placed in this group. Phage ϕ PV83 is the only Sa5int member which encodes a known virulence factor (*lukM*). Finally, serogroups A, B, and Fa and two holin genes, 303 and 438, were found in phage group Sa6int. These phages typically integrate into the lipase gene (*geh*) of *S. aureus* (21).

A lower degree of variation regarding the three analyzed modules was found in the *S. aureus* phage groups Sa1int, Sa2int, Sa4int, and Sa7int (Fig. 1). Some Sa1int phages harbor the exfoliative toxin a (*eta*). Sa2int phages integrate into an ORF (SA3121) of unknown function in the *S. aureus* genome, as shown for phage ϕ 12 of strain 8325 and the PVL-encoding phages (18). All Sa7int phages contained the serogroup B module and holin gene 303 or 438, with the exception of prophage ϕ 6390. This phage encodes holin 255a, which is characteristic for Sa3int phages, the typical *sak*-carrying phages. We recently determined the integration site of ϕ 6390 (intergenic region between *rpmF* and *isdB*) in the prototypic *S. aureus* strain RN6390 and showed that it carries the virulence gene *sak* (11).

The remaining *S. aureus* phages harbored unique integrase types but shared serogroups or holin genes with other integrase groups. For instance, the prophage ϕ RF122 of the bovine *S. aureus* isolate RF122 (Sa8int) and the Sa9int and Sa12int phages all exhibit the serogroup B module and holin type 438 or 303. None of the *S. aureus* bacteriophage modules could be detected in the phages of the other staphylococcal species.

The results give an overview of phage mosaicism, but they also show strong association between functional units. Especially, the phage-encoded virulence factors were closely linked to the integrase groups.

Identification of *S. aureus* prophages by multiplex PCR. For a prevalence analysis of *S. aureus* prophages in a large strain collection, we focused on the identification of the integrase polymorphism for several reasons. First, nucleotide sequences

FIG. 2. Multiplex PCR detecting the Sa1int to Sa7int integrase genes in prototypic *S. aureus* strains.

are well conserved within integrase groups making, the gene an ideal target for PCR amplification. Second, the integrase-defined grouping had the best discriminatory power, reflecting the diversity of the *S. aureus* phage population as well as their relatedness. Last, the integrase type is closely linked to the virulence gene content of the prophage and might therefore convey information about the *S. aureus* pathogenic potential (Fig. 1).

We established a multiplex PCR scheme to discriminate between the seven most prominent *S. aureus* integrase families, Sa1int to Sa7int. The method reliably identified the prophage content of the prototypic *S. aureus* strains N315, Mu50, MW2, MRSA252, MSSA476, Newman, and 8325 (Fig. 2). Additionally the method was validated with *S. aureus* strains 8325-4 and 1039 lysogenized with genome-sequenced phages of the International Typing Set. The seven *int* groups were detectable by the multiplex PCR in these isolates (data not shown).

Distribution of phage types in *S. aureus* clonal lineages. To acquire a representative collection of *S. aureus* strains from different clonal lineages, 386 isolates were obtained from the following sources: 161 isolates from nasal colonization of healthy individuals, 115 blood culture isolates, 73 isolates from diverse clinical samples, and 37 reference strains. Both MRSA and methicillin-susceptible *S. aureus* strains were included in this collection. Isolates were typed by either *spa* or PFGE typing and assigned to MLST CCs. After excluding singletons and CCs with fewer than 10 isolates from the strain collection, 291 isolates remained for further analyses. These strains were grouped into seven CCs representing different *agr* types: CC5 (*agr*-2), CC8 (*agr*-1), CC15 (*agr*-2), CC22 (*agr*-1), CC25 (*agr*-1), CC30 (*agr*-3), and CC45 (*agr*-1). No substantial difference in the distribution of CCs was observed in the distinct geographical locations.

Analyzing the prevalence of the seven phage groups in the 291 *S. aureus* isolates revealed that prophages of the groups Sa3int (74%) and Sa2int (33%) (Table 1) were the most frequent ones. Sa7int could be detected in 16%, Sa1int in 9%, Sa6int in 6%, and Sa4int only in 0.2% of the isolates; 13% of the isolates harbored none of the targeted prophages. Next we

FIG. 1. In silico analysis of the integrase groups, serogroups, holin groups, and virulence genes of 70 published staphylococcal bacteriophages of the class *Siphoviridae*. Integrase nucleotide sequences were aligned using the ClustalW algorithm. Identical serogroups and holin groups are color coded. Integrases of the serine recombinase-type family are shaded in gray.

TABLE 1. Distribution of the seven prophage groups Sa1int to Sa7int in common *S. aureus* CCs

CC	% (no.) of isolates	% (no. of isolates) ^a with prophage group:							
		Sa1int	Sa2int	Sa3int	Sa4int	Sa5int	Sa6int	Sa7int	None ^d
15	9 (32)	9 (3)	44 (14)	0 (0) ^{***b}	0 (0)	3 (1)	19 (6)	6 (2)	31 (10) ^{*c}
22	4 (13)	8 (1)	23 (3)	85 (11)	0 (0)	0 (0)	0 (0)	0 (0)	15 (2)
25	9 (30)	0 (0)	47 (14)	86 (26)	0 (0)	0 (0)	0 (0)	0 (0) ^{*b}	10 (3)
30	20 (67)	1 (1) ^{*b}	58 (39) ^{***c}	85 (57)	0 (0)	33 (22) ^{***c}	10 (7)	16 (11)	6 (4)
45	15 (50)	26 (13) ^{***c}	4 (2) ^{***b}	72 (36)	0 (0)	18 (9)	2 (1)	14 (7)	24 (12)
5	18 (63)	11 (7)	17 (11) ^{*b}	89 (56)	2 (1)	5 (3)	5 (3)	38 (24) ^{***c}	5 (3)
8	11 (36)	6 (2)	39 (14)	83 (30)	0 (0)	3 (1)	0 (0)	11 (4)	11 (4)
All	100 (291)	9 (27)	33 (97)	74 (216)	0.2 (1)	12 (36)	6 (17)	16 (48)	13 (38)

^a ***, $P < 0.001$; **, $P < 0.005$; *, $P < 0.05$.^b Significantly lower than in the whole cohort.^c Significantly higher than in the whole cohort.^d Absence of any of the seven prophage groups.

asked whether there was a relationship between prophage groups and *S. aureus* clonal lineages. The likelihood ratio test was applied to compare the frequency of phage types in certain CCs with that in the whole cohort. Indeed, the frequency of the different phage types varied remarkably between the tested *S. aureus* CCs. In CC15, the very common Sa3int phages were never detected ($P < 0.001$). This result was verified by Southern analysis using probes specific for the phage-encoded staphylokinase (*sak*) (data not shown). Additionally, in significantly more ($P < 0.05$) CC15 isolates, none of the seven prophage groups could be detected. No CC25 isolate harbored a Sa7int phage ($P < 0.05$). In CC30 isolates, Sa1int phages were significantly less frequent ($P < 0.05$), but the Sa2int and Sa5int phages were significantly more frequent ($P < 0.001$ for both) than in the whole *S. aureus* strain collection. In CC45 the prevalence of Sa1int phages was significantly higher ($P < 0.005$) and that of Sa2int phages significantly lower ($P < 0.001$) than in all isolates. In CC5 also the Sa2int group was less often detected ($P < 0.05$), whereas Sa7int phages were significantly more frequent ($P < 0.001$). The phage prevalences in CC22 and CC8 did not differ from those in the overall cohort.

We next calculated the number of simultaneously occurring prophages per bacterial cell. We could detect none of the seven prophages in 13% (38/291) of the isolates, one prophage in 36% (106/291), two in 38% (110/219), three in 11% (32/219), and four in 2% (4/219). Thus, most isolates contain one or two prophages, but none contain more than four. By linking these results with the genetic background of the isolates, we could show that in CC15, strains with no or only one prophage were strongly overrepresented ($P < 0.0004$), whereas CC30 isolates were more often than average lysogenic for at least two prophages ($P < 0.0001$).

Distribution of phages in invasive versus colonizing isolates.

To test whether phage prevalences differ in invasive and colonizing *S. aureus* populations, blood culture isolates were compared to nasal carriage strains. In total 276 isolates were available for analysis, 115 from blood cultures and 161 nasal isolates. When applying the multiplex PCR scheme, we could show that in the colonizing population significantly more isolates harbor Sa3int phages than in the invasive strains ($P < 0.05$) (Table 2). No differences were observed in the prevalence of the other phage types. Most of the isolates were lysogenic for one or more phages: in only 21% of the blood

culture and 13% of the nasal isolates could none of the seven phage groups be detected. Additionally, the isolates from both populations also did not differ in the number of prophages per cell (data not shown).

The higher frequency of Sa3int in nasal isolates was not correlated with an overrepresentation of certain CCs (data not shown). In general, no CC was linked to invasive or colonizing strains, and phage distribution was associated only with the genetic background of the strain and not with its origin.

DISCUSSION

Bacteriophages have a tremendous impact on the biology of their bacterial hosts, because they play an important role in bacterial ecology, evolution, and adaptation. For instance, in the human pathogen *S. aureus*, prophages are responsible for the emergence and evolution of new threatening strains such as the community-acquired MRSA strains which carry PVL-encoding prophages. Despite their importance, a comprehensive picture of the distribution of prophages in the *S. aureus* strain populations was lacking. In the present study we could show that prophage prevalence was associated with the clonal background of *S. aureus*, indicating that the spread of the phages in the bacterial population is at least partially restricted. In certain CCs some phage groups were completely absent and others were significantly less or, on the other hand, significantly more frequent. The most prominent disequilib-

TABLE 2. Distribution of the seven prophage groups Sa1int to Sa7int in nasal carriage and blood culture populations

Prophage group	% (no.) of isolates ^a from:		
	Blood culture (<i>n</i> = 115)	Nose (<i>n</i> = 161)	Total (<i>n</i> = 276)
Sa1int	11 (13)	8 (13)	9 (26)
Sa2int	27 (31)	32 (51)	30 (82)
Sa3int	59 (68)	73 (118) [*]	67 (186)
Sa4int	0 (0)	0 (0)	0 (0)
Sa5int	9 (10)	17 (27)	13 (37)
Sa6int	6 (7)	9 (14)	8 (21)
Sa7int	13 (15)	9 (15)	11 (30)
None	21 (24)	13 (21)	16 (45)

^a *, $P < 0.05$.

rium was the finding that CC15 strains do not carry Sa3int phages, although this is the most common phage group found in *S. aureus*, with a prevalence of up to 90% (11, 25, 33). In good agreement with this total absence of Sa3int phages from CC15 strains is the earlier observation that this lineage completely lacks staphylococcal superantigen genes (15), some of which (i.e., *sea*, *sep*, *sek*, and *seq*) are carried on Sa3int phages. In addition, many isolates from the CC15 complex carried none of the seven prophage groups, suggesting that this lineage is particularly restrictive to the uptake of foreign DNA. In CC30 an unusually high proportion of Sa2int phages could be detected. In this *S. aureus* lineage an early pandemic clone, which already carried the PVL-encoding phage (members of the Sa2int group), developed into a community-acquired methicillin-resistant clone by acquiring the SCCmec type IV cassette (30). CC5 isolates were characterized by the high proportion of Sa7int phages. Interestingly, Sa7int phages of this CC often carry the *sak* gene detected previously on phage ϕ 6390 (11) and phages of selected clinical isolates (unpublished data).

The CCs analyzed were shown to differ in their R-M specificity genes (34). The SauI R-M system is a major barrier to horizontal gene transfer in *S. aureus* and seems to delay the evolution of new strains. Mobile genetic elements present in one strain will move to a strain of the same lineage at a higher frequency than to strains of other lineages. As a consequence, *S. aureus* lineages carry a unique combination of core variable genes, suggesting only a vertical transmission of these genes (24). Additional R-M systems were described for *S. aureus*, some of which were shown to be phage encoded, which may also contribute to phage exclusion (7). In addition to the host restriction, the lysogenic immunity of a resident prophage may play a role in prevalence differences. Indeed, Sa1int- and Sa2int-type phages appear to be (in part) mutually exclusive; the simultaneous occurrence of both in a single isolate is uncommon ($P = 0.0095$). In concordance, in CC30, Sa1int phages were rare whereas Sa2int phages were frequent; in CC45, the distribution is vice versa.

When comparing the patterns of phage prevalence in invasive versus colonizing *S. aureus* isolates, no differences were detected with the exception of Sa3int phages, which were significantly more common in colonizing strains. This is in agreement with our own observation that in 96% of nasal isolates Sa3 phages were stably integrated into the *hly* gene (11). Hly-converting phages encode the immune-modulatory proteins Sak, Scin, and Chips (6, 33). These may act together to resist the innate immune response encountered during nasal colonization (complement, defensins, and phagocytosis). In contrast, the lack of the Sa3int phages in infecting isolates is correlated to restore Hly production. This is in concordance with previous findings indicating that Hly-producing strains are linked to infectious conditions (11, 16, 29). However, when comparing community-acquired invasive isolates with nasal carriage isolates, Lindsay et al. were unable to detect any association between gene and invasive isolates (24). This discrepancy is perhaps due to different criteria for the inclusion of isolates in the invasive group.

In this study we developed a reliable classification scheme for staphylococcal phages of the *Siphoviridae* family, which is the largest, best-described group of temperate *S. aureus*

phages. We could show that phages can be clustered into defined groups based on the integrase sequence. This feature fulfils the criteria to be discriminative enough to account for the high diversity of the prophages without being too diverse, thus creating only types represented by single members. Importantly, the integrase identification allows prediction of the chromosomal location of the prophage and gives an indication of the virulence gene content. Analysis of a large *S. aureus* strain collection revealed that most of the isolates contained one to three prophages, which is in line with the phage content of the *S. aureus* strains for which the whole genomes have been sequenced. The most prevalent phages were the Sa3int group, followed by Sa2int. Sa4int was detected only once.

To assess phage diversity, 71 complete staphylococcal bacteriophage genome sequences from the databases were analyzed in three marker regions: lysogeny module, morphogeny module, and lysis module. Various degrees of genome mosaicism could be observed within the different Sa-int groups. The Sa1int and Sa2int groups were characterized by a uniform modular architecture with strong links between the genes for integrase, holin, and encoded virulence factors. Perhaps it is evolutionarily beneficial to interchange this whole unit, which is in proximity in the circular form of the phage. Multiple alignments of several PVL-encoding phages revealed a high degree of mosaic structure of the phage genomes, but the *luk*-PV genes were always located in a 6.4-kb region consisting of the host lysis module, *luk*-PV, *attP*, and the integrase gene (18). We aligned the eight PVL-carrying and the three non-PVL-carrying strains of the Sa2int group to discern the cross-over point for integration of the toxin complex. This point appeared to be located at the end of the phage amidase ORF (data not shown). The close organization of the lytic module and the inserted virulence factors is perhaps favored to optimize the phage control of the expression of the pathogenicity genes (32). Interestingly, Sa1int and Sa2int phages, which excluded one another, did not share any modules, suggesting a parallel evolution with no or little contact. It would now be of interest to test whether recombination occurs only within the lineage boundaries or whether phage mosaics are evolutionarily more ancient than the lineage branching. The fact that none of the *S. aureus* phage modules are present in phages from other staphylococcal species argues in favor of the first possibility.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft through grants to C.W. (Wo578/6 and TR34), B.M.B. (TR34 and GRK-840), and D.G. (GRK-840). The contributions of R.P. and J.D. were supported by grant LSHM-CT-2006-019064 from the European Union. Isolates NRS158, NRS161, NRS162, NRS184, NRS187, NRS192, NRS226, NRS229, NRS232, NRS237, and NRS384 were obtained through the Network on Antimicrobial Resistance in *Staphylococcus aureus* (NARSA) program supported under NIAID/NIH contract no. HHSN2722007 00055C.

REFERENCES

- Ackermann, H. W., and M. S. DuBow. 1987. Natural groups of bacteriophages, viruses of prokaryotes, vol. 2. CRC Press, Boca Raton, FL.
- Baba, T., F. Takeuchi, M. Kuroda, H. Yuzawa, K. Aoki, A. Oguchi, Y. Nagai, N. Iwama, K. Asano, T. Naimi, H. Kuroda, L. Cui, K. Yamamoto, and K. Hiramatsu. 2002. Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* **359**:1819–1827.
- Betley, M. J., and J. J. Mekalanos. 1985. Staphylococcal enterotoxin A is encoded by phage. *Science* **229**:185–187.

4. Brussow, H., and F. Desiere. 2001. Comparative phage genomics and the evolution of *Siphoviridae*: insights from dairy phages. *Mol. Microbiol.* **39**: 213–222.
5. Canchaya, C., C. Proux, G. Fournous, A. Bruttin, and H. Brussow. 2003. Prophage genomics. *Microbiol. Mol. Biol. Rev.* **67**:238–276.
6. Coleman, D. C., D. J. Sullivan, R. J. Russell, J. P. Arbutnot, B. F. Carey, and H. M. Pomeroy. 1989. *Staphylococcus aureus* bacteriophages mediating the simultaneous lysogenic conversion of beta-lysin, staphylokinase and enterotoxin A: molecular mechanism of triple conversion. *J. Gen. Microbiol.* **135**:1679–1697.
7. Dempsey, R. M., D. Carroll, H. Kong, L. Higgins, C. T. Keane, and D. C. Coleman. 2005. Sau42I, a BglI-like restriction-modification system encoded by the *Staphylococcus aureus* quadruple-converting phage Phi42. *Microbiology* **151**:1301–1311.
8. Enright, M. C., N. P. Day, C. E. Davies, S. J. Peacock, and B. G. Spratt. 2000. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J. Clin. Microbiol.* **38**:1008–1015.
9. Feil, E. J., J. E. Cooper, H. Grundmann, D. A. Robinson, M. C. Enright, T. Berendt, S. J. Peacock, J. M. Smith, M. Murphy, B. G. Spratt, C. E. Moore, and N. P. Day. 2003. How clonal is *Staphylococcus aureus*? *J. Bacteriol.* **185**:3307–3316.
10. Goerke, C., S. Matias y Papenberg, S. Dasbach, K. Dietz, R. Ziebach, B. C. Kahl, and C. Wolz. 2004. Increased frequency of genomic alterations in *Staphylococcus aureus* during chronic infection is in part due to phage mobilization. *J. Infect. Dis.* **189**:724–734.
11. Goerke, C., C. Wirtz, U. Fluckiger, and C. Wolz. 2006. Extensive phage dynamics in *Staphylococcus aureus* contributes to adaptation to the human host during infection. *Mol. Microbiol.* **61**:1673–1685.
12. Goerke, C., M. Gressinger, K. Endler, C. Breitkopf, K. Wardecki, M. Stern, C. Wolz, and B. C. Kahl. 2007. High phenotypic diversity in infecting but not in colonizing *Staphylococcus aureus* populations. *Environ. Microbiol.* **9**:3134–3142.
13. Harmsen, D., H. Claus, W. Witte, J. Rothganger, D. Turnwald, and U. Vogel. 2003. Typing of methicillin-resistant *Staphylococcus aureus* in a university hospital setting by using novel software for *spa* repeat determination and database management. *J. Clin. Microbiol.* **41**:5442–5448.
14. Holden, M. T., E. J. Feil, J. A. Lindsay, S. J. Peacock, N. P. Day, M. C. Enright, T. J. Foster, C. E. Moore, L. Hurst, R. Atkin, A. Barron, N. Bason, S. D. Bentley, C. Chillingworth, T. Chillingworth, C. Churcher, L. Clark, C. Corton, A. Cronin, J. Doggett, L. Dowd, T. Felwell, Z. Hance, B. Harris, H. Hauser, S. Holroyd, K. Jagels, K. D. James, N. Lennard, A. Line, R. Mayes, S. Moule, K. Mungall, D. Ormond, M. A. Quail, E. Rabinowitsch, K. Rutherford, M. Sanders, S. Sharp, M. Simmonds, K. Stevens, S. Whitehead, B. G. Barrell, B. G. Spratt, and J. Parkhill. 2004. Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. *Proc. Natl. Acad. Sci. USA* **101**:9786–9791.
15. Holtfreter, S., D. Grumann, M. Schmudde, H. T. Nguyen, P. Eichler, B. Strommenger, K. Kopron, J. Kolata, S. Giedrys-Kalembe, I. Steinmetz, W. Witte, and B. M. Broker. 2007. Clonal distribution of superantigen genes in clinical *Staphylococcus aureus* isolates. *J. Clin. Microbiol.* **45**:2669–2680.
16. Jin, T., M. Bokarewa, L. McIntyre, A. Tarkowski, G. R. Corey, L. B. Reller, and V. G. Fowler, Jr. 2003. Fatal outcome of bacteraemic patients caused by infection with staphylokinase-deficient *Staphylococcus aureus* strains. *J. Med. Microbiol.* **52**:919–923.
17. Kaneko, J., T. Kimura, S. Narita, T. Tomita, and Y. Kamio. 1998. Complete nucleotide sequence and molecular characterization of the temperate staphylococcal bacteriophage phiPVL carrying Pantone-Valentine leukocidin genes. *Gene* **215**:57–67.
18. Kaneko, J., and Y. Kamio. 2004. Bacterial two-component and hetero-heptameric pore-forming cytolytic toxins: structures, pore-forming mechanism, and organization of the genes. *Biosci. Biotechnol. Biochem.* **68**:981–1003.
19. Kuroda, M., T. Ohta, I. Uchiyama, T. Baba, H. Yuzawa, I. Kobayashi, L. Cui, A. Oguchi, K. Aoki, Y. Nagai, J. Lian, T. Ito, M. Kanamori, H. Matsumaru, A. Maruyama, H. Murakami, A. Hosoyama, Y. Mizutani-Ui, N. K. Takahashi, T. Sawano, R. Inoue, C. Kaito, K. Sekimizu, H. Hirakawa, S. Kuhara, S. Goto, J. Yabuzaki, M. Kanehisa, A. Yamashita, K. Oshima, K. Furuya, C. Yoshino, T. Shiba, M. Hattori, N. Ogasawara, H. Hayashi, and K. Hiramatsu. 2001. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet* **357**:1225–1240.
20. Kwan, T., J. Liu, M. DuBow, P. Gros, and J. Pelletier. 2005. The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. *Proc. Natl. Acad. Sci. USA* **102**:5174–5179.
21. Lee, C. Y., and S. L. Buranen. 1989. Extent of the DNA sequence required in integration of staphylococcal bacteriophage L54a. *J. Bacteriol.* **171**:1652–1657.
22. Lindsay, J. 2008. *Staphylococcus: molecular genetics*. Caister Academics, Norfolk, United Kingdom.
23. Lindsay, J. A., and M. T. Holden. 2004. *Staphylococcus aureus*: superbug, super genome? *Trends Microbiol.* **12**:378–385.
24. Lindsay, J. A., C. E. Moore, N. P. Day, S. J. Peacock, A. A. Witney, R. A. Stabler, S. E. Husain, P. D. Butcher, and J. Hinds. 2006. Microarrays reveal that each of the ten dominant lineages of *Staphylococcus aureus* has a unique combination of surface-associated and regulatory genes. *J. Bacteriol.* **188**:669–676.
25. Matthews, A. M., and R. Novick. 2005. Staphylococcal phages, p. 297–318. In M. K. Waldor, D. I. Friedman, and S. L. Adhya (ed.), *Phages. Their role in bacterial pathogenesis and biotechnology*. ASM Press, Washington, DC.
26. Narita, S., J. Kaneko, J. Chiba, Y. Piemont, S. Jarraud, J. Etienne, and Y. Kamio. 2001. Phage conversion of Pantone-Valentine leukocidin in *Staphylococcus aureus*: molecular analysis of a PVL-converting phage, phiSLT. *Gene* **268**:195–206.
27. Novick, R. P. 2003. Mobile genetic elements and bacterial toxinoses: the superantigen-encoding pathogenicity islands of *Staphylococcus aureus*. *Plasmid* **49**:93–105.
28. Pantucek, R., J. Doskar, V. Ruzickova, P. Kasperek, E. Oracova, V. Kvardova, and S. Rosypal. 2004. Identification of bacteriophage types and their carriage in *Staphylococcus aureus*. *Arch. Virol.* **149**:1689–1703.
29. Peacock, S. J., C. E. Moore, A. Justice, M. Kantzanou, L. Story, K. Mackie, G. O'Neill, and N. P. Day. 2002. Virulent combinations of adhesin and toxin genes in natural populations of *Staphylococcus aureus*. *Infect. Immun.* **70**:4987–4996.
30. Robinson, D. A., A. M. Kearns, A. Holmes, D. Morrison, H. Grundmann, G. Edwards, F. G. O'Brien, F. C. Tenover, L. K. McDougal, A. B. Monk, and M. C. Enright. 2005. Re-emergence of early pandemic *Staphylococcus aureus* as a community-acquired methicillin-resistant clone. *Lancet* **365**:1256–1258.
31. Strommenger, B., C. Kettlitz, T. Weniger, D. Harmsen, A. W. Friedrich, and W. Witte. 2006. Assignment of *Staphylococcus* isolates to groups by *spa* typing, *Sma*I macrorestriction analysis, and multilocus sequence typing. *J. Clin. Microbiol.* **44**:2533–2540.
32. Sumby, P., and M. K. Waldor. 2003. Transcription of the toxin genes present within the *Staphylococcal* phage phiSa3ms is intimately linked with the phage's life cycle. *J. Bacteriol.* **185**:6841–6851.
33. van Wamel, W. J., S. H. Rooijakkers, M. Ruyken, K. P. van Kessel, and J. A. van Strijp. 2006. The innate immune modulators staphylococcal complement inhibitor and chemotaxis inhibitory protein of *Staphylococcus aureus* are located on beta-hemolysin-converting bacteriophages. *J. Bacteriol.* **188**:1310–1315.
34. Waldron, D. E., and J. A. Lindsay. 2006. Saul: 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**:5578–5585.
35. Yamaguchi, T., T. Hayashi, H. Takami, K. Nakasone, M. Ohnishi, K. Nakayama, S. Yamada, H. Komatsuzawa, and M. Sugai. 2000. Phage conversion of exfoliative toxin A production in *Staphylococcus aureus*. *Mol. Microbiol.* **38**:694–705.