Genetic Variation among Pantone-Valentine Leukocidin-Encoding Bacteriophages in *Staphylococcus aureus* Clonal Complex 30 Strains

Liang Chen,^a Kalyan D. Chavda,^a Mihir Solanki,^a José R. Mediavilla,^a Barun Mathema,^a Patrick M. Schlievert,^b Barry N. Kreiswirth^a

Public Health Research Institute Center, International Center for Public Health, University of Medicine and Dentistry of New Jersey, Newark, New Jersey, USA^a; Department of Microbiology, Carver College of Medicine, Iowa City, Iowa, USA^b

Clonal complex 30 (CC30), one of the major *Staphylococcus aureus* lineages, has caused extensive hospital-acquired and community-acquired infections worldwide. Recent comparative genomics studies have demonstrated that three CC30 clones—phage type 80/81, Southwest Pacific (SWP), and contemporary EMRSA-16 associated (Con) strains—shared a recent common ancestor more than 100 years ago. Pantone-Valentine leukocidin (PVL), a bacteriophage encoded toxin that has been epidemiologically linked with community-associated methicillin-resistant *S. aureus* (CA-MRSA), has frequently been identified in CC30 clones, although the *pvl* gene variation and distribution of PVL-encoding phages are poorly understood. We determined here the distribution of PVL phages, PVL gene sequences, and chromosomal phage insertion sites in 52 *S. aureus* CC30 PVL-harboring isolates, collected from four continents over a 75-year period. Our results indicate that PVL phages with icosahedral heads, including Φ 108PVL and Φ PVL, were mainly associated with phage 80/81 strains, whereas phages with elongated heads were predominantly found in SWP (Φ Sa2958 and Φ TCH60) and Con (Φ Sa2USA) strains. Nine single-nucleotide polymorphisms were identified in the *lukSF-PV* gene, with six isolates harboring the R variant that has been previously associated with CA-MRSA strains. Interestingly, all six R variant strains belonged to the same Con CC30 clone and carried a Φ Sa2USA-like phage. Similar chromosomal phage insertion sites were also identified in all 52 PVL-harboring CC30 strains. These analyses provide important insights into the microepidemiology of PVL-harboring CC30 strains, while the discovery of Φ Sa2USA-associated R variant strains sheds further light on the evolution of PVL-positive CA-MRSA.

During the last 60 years, *Staphylococcus aureus* clonal complex 30 (CC30) has had a profound impact on global human health, as evidenced by three pandemic “waves,” as well as the toxic shock syndrome (TSS) epidemic (1–3). Three CC30 clones have caused a significant worldwide disease burden in both community and health care settings, including the historic phage-type 80/81 strain (80/81), the contemporary community-associated methicillin-resistant *S. aureus* (CA-MRSA) Southwest Pacific (SWP) clone, and the contemporary hospital-associated MRSA (HA-MRSA) clone EMRSA-16 and its related methicillin-susceptible *S. aureus* (MSSA) epidemic strains (Con) (1–3). The evolutionary relatedness of these three clones has recently been deciphered (2, 4). Comparative genome studies have demonstrated that all three clones evolved from a recent common ancestor that existed over 100 years ago (2, 4), and the timing of emergence for each clone has been estimated as 1936 (phage 80/81), 1967 (SWP), and 1975 (EMRSA-16), respectively (2).

Notably, both the 80/81 and SWP clones are associated with the production of Pantone-Valentine leukocidin (PVL) (1), a bacteriophage-encoded, bicomponent, hetero-oligomeric, pore-forming cytolytic toxin initially described in 1932 (5). The genes encoding PVL (*lukS-PV* and *lukF-PV*) have been identified in different temperate bacteriophages bearing both icosahedral and elongated capsids (heads), including Φ PVL (6), Φ 108PVL (7), Φ 7247PVL (8), Φ Sa2958 (9), Φ Sa2MW (10), Φ SLT (11), Φ Sa2USA (12), Φ TCH60 (GenBank accession no. CP002110), and Φ tp310-1 (EF462197). Sequence variations within the *lukSF-PV* gene, as well as several PVL protein isoforms bearing nonsynonymous mutations, have recently been described (13–19), although their functional implications are still not clear. Several studies have explored the distribution of PVL bacteriophages and gene variants in different *S. aureus* clones (14, 20, 21), in

particular the CA-MRSA CC8 clone USA300. In contrast, PVL bacteriophages and gene variants in CC30 strains, along with the distribution of these within different pandemic lineages, have not been addressed. We examined here a collection of 52 PVL-harboring CC30 *S. aureus* strains, collected from four continents over a period of 75 years, in order to shed light on the evolution of PVL within this significant lineage.

MATERIALS AND METHODS

Bacterial strains. A total of 34 PVL-harboring CC30 *S. aureus* strains were selected from our previous study (4), and an additional 18 strains were included to maximize spatial and temporal diversity (Table 1). Forty of the strains were obtained from either the Network on Antimicrobial Resistance in *Staphylococcus aureus* (NARSA; <http://www.narsa.net>) or from our own collection at the Public Health Research Institute Tuberculosis Center. Eleven isolates were kindly provided by D. A. Robinson (University of Mississippi Medical Center), and strain 80/108 was a gift from K. Hiramatsu (Juntendo University).

Identification of CC30 clones. Genomic DNA was extracted using a Wizard genomic DNA purification kit (Promega, Madison, WI), as previously described (22). All isolates were characterized by multilocus sequence typing (MLST) (23), *spa* typing (24), SCCmec typing (25), and SCCmec IV subtyping (26). In our previous study, a number of CC30-specific (e.g., phage 80/81, SWP, and Con) single nucleotide polymor-

Received 14 November 2012 Returned for modification 13 December 2012
Accepted 27 December 2012

Published ahead of print 2 January 2013

Address correspondence to Barry N. Kreiswirth, kreiswba@umdnj.edu.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JCM.03015-12

TABLE 1 Characteristics of 52 PVL-positive CC30 *S. aureus* isolates

Strain	Alias	Yr	Location	MLST ^a	spa type ^b	SCCmec ^c	agrC ^d	hla ^d	Lineage	arR cluster ^e	PVL		Nucleotide (reference sequence, Φ 108PVL) ^f																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
											phage head ^f	PVL phage ^f	PVL haplotype	lukS-PV (nt 1 to 939)											lukF-PV (nt 941 to 1918)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
														(C)	(T)	(A)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)	(G)	(G)	(C)

^a The allelic profiles for each MLST sequence type (ST) are as follows (all belong to CC30): ST30(2-2-2-2-6-3-2), ST37(2-2-2-2-15-3-2), ST43(2-2-2-2-5-2-6-3-2), ST431(2-2-2-2-54-6-3-2), ST1472(2-2-2-2-6-3-162), ST2385(259-2-2-2-6-3-2), and ST2386(2-342-2-2-6-3-2).

^b *spa* types are denoted by both eGenomics and Ridom *spa* type numbers. n1, novel eGenomics *spa* motif WKKAKAO; n2, novel eGenomics *spa* motif XAKAOMQ.

^c “—” indicates MSSA (no SCCmec elements detected).

^d Single nucleotide polymorphisms (SNPs) in the *hla* and *agrC* genes used to classify different CC30 sublineages: *hla* (G/A substitution at position 1181065 in MRS252) and *agrC* (G/A substitution at position 2184961 in MRS252).

^e The *arR* sequences of cluster A strains are identical to those of strain 81/108, whereas the *arR* sequences of cluster B are the same as those of strain TCH60; the *arR* sequences of 15071 (cluster C) and 23417 (cluster D) have one and two nucleotide differences relative to those of cluster A and cluster B, respectively.

^f *i*, icosahedral head phage type; E, elongated head phage type; NT, nontypable.

^g Nucleotides differing from the reference sequence are underlined.

^h Strain 21301, characterized by a mutant SNP in *agrC* (A), but a wild-type SNP in *hla* (G), was classified within the 80/81 group, based on other lineage specific SNPs including *isdH* and *lssS* (data not shown) (4).

phisms (SNPs) were identified (4). Among these, the premature stop codon in the *hla* gene (G to A at position 1181065 in MRSA252, referred to hereafter as *hla*_{STOP}), and the nonsynonymous mutation in *agrC* (G to A at position 2184961 in MRSA252, referred to hereafter as *agrC*_{G55R}), can be used to differentiate Con strains from phage 80/81 and SWP strains (2, 4). Identification of *hla*_{STOP} and *agrC*_{G55R} were performed using a method previously described (4). Based on the presence of these two SNPs, and of the SCCmec IV element, the 52 CC30 isolates were assigned into three groups (phage 80/81, SWP, and Con) (Table 1). Phage 80/81 and SWP strains are wild type (G) for both the *hla* and the *agrC* SNPs, while SWP strains can be further differentiated from 80/81 strains by carriage of the SCCmec IV element. In contrast, CC30 Con strains are mutated (A) at both SNP positions.

PVL phage and gene identification. The full-length *lukSF-PV* gene (1,918 bp) was amplified by PCR and sequenced according to previously published methods (16). Additional PCRs spanning the proximal and distal junctions of the phage-chromosome DNA junctions were performed according to the method described by Boakes et al. (16). The PCR products were sequenced in both directions using standard Sanger DNA sequencing methods.

A previously published eight-reaction PCR assay was performed to detect five of the PVL-encoding phages (ΦSa2958, ΦSa2mw, ΦPVL, Φ108PVL, and ΦSLT) (9). Among these, PCR-1 and PCR-2 were designed to identify carriage of two morphologically distinct phages by targeting genes encoding an icosahedral versus an elongated head shape. PCR-3 and PCR-4 were designed to link these morphologically specific tail genes to the PVL genes. The remaining reactions (PCR-5 to -8) were used to classify individual PVL phages, including ΦPVL, Φ108PVL, ΦSa2958, ΦSa2MW, and ΦSLT (9). Two additional PCRs were included in the present study in order to further identify ΦSa2USA (16) (PCR-9) and Φ7424PVL (8) (PCR-10, primer sets A and B), using previously published methods. PCR-9 targets the region between *lukS-PV* and the neighboring ΦSLT ORF484-like gene (SAUSA300_1383) within ΦSa2USA (16). PCR-10 includes two sets of primers: primer set A targets the region between *int* (integrase) and *rep* (repressor protein); while primer set B targets the region between *lukS-PV* and *mtp* (phage tail tape-measure protein) within Φ7424PVL (8). A new primer set (PCR-11) was also designed in order to detect ΦTCH60, using primers intF2 (ATGTT TTCGAGT TTTT GAGTTAG) (9) and TCH60aR (TAACAACGATCA ACAACCA) to generate an amplicon 2,675 bp in length, targeting the region between *int* and a hypothetical protein coding sequence (HMPREF0772_11656). Examining the sequence of Φtp310-1 (GenBank accession no. EF462197) revealed that it harbors an icosahedral head (positive for PCR-1 and PCR-3) and can be identified by the presence of PCR products specific for ΦPVL and Φ108PVL. Therefore, a total of nine PVL-encoding phages (ΦSa2958, ΦSa2mw, ΦPVL, Φ108PVL, ΦSLT, ΦSa2USA, Φ7424PVL, ΦTCH60, and Φtp310-1) can be identified using 11 sets of PCR assays.

Phage that were positive by characterization assays (PCR-5 to -11) and icosahedral/elongated head type assays (PCR-1 and PCR-2), but which were negative for PCR spanning morphologically specific tail genes to the PVL genes (PCR-3 and -4), were described as “phage-like” based on the presence of known individual phage types (e.g., Φ108PV-like and ΦPVL-like phage, Table 1). Phages that were not identifiable by the characterization assays (PCR-5 to -11) were described as “NT” (nontypeable).

Sequence analysis. DNA sequences of *lukSF-PV* genes, as well as the proximal and distal junctions of phage chromosomal insertion sites, were aligned using CLUSTAL X2.0 (27). Statistical parsimony (haplotype) networks were generated by TCS (version 1.21) (28) using the method described previously (14). For the junction regions, MEGA5.1 (29) was used to construct a maximum-likelihood phylogenetic tree using the Hasegawa-Kishino-Yano model (30) with 500 bootstrap replicates.

Statistical analysis. The χ^2 test (or the Fisher exact test) was used to compare the prevalence of phages in different CC30 clones. *P* values of ≤ 0.05 were considered significant.

RESULTS

Characterization of *S. aureus* CC30 strains. A total of 52 strains were collected from different geographical regions, including North America, Europe, Asia, and Oceania, spanning a period of 75 years (1935 to 2010). MLST and *spa* typing demonstrated that all 52 isolates belonged to CC30 (Table 1). Based on the presence of SCCmec IV and two lineage-specific SNPs (*hla*_{STOP} and *agrC*_{G55R}), these isolates were further classified into three lineages, i.e., 80/81 (*n* = 31), SWP (*n* = 14), and Con (*n* = 7). Strain 21301, characterized by a mutant SNP in *agrC* (A), but a wild-type SNP in *hla* (G), was nevertheless classified within the 80/81 group, based on other lineage-specific SNPs within *isdH* and *hssS* (data not shown) (4).

Distribution of PVL phages. Using the PCR methods described above, all 52 strains were successfully assigned into either icosahedral or elongated head morphology groups (Table 1). Although PVL phages belonging to both groups were identified in all three CC30 sublineages, the relative frequency varied by lineage. Phage type 80/81 strains frequently carry icosahedral head phages (93.5%) (*P* < 0.001 versus SWP and Con), while SWP and Con strains mainly harbor elongated head phages (92.8 and 85.7% for SWP and Con, respectively; *P* < 0.001 versus phage 80/81).

Distinct PVL phages were also identified within CC30 strains. Within the 80/81 sublineage, most of the isolates tested harbored one of the following known PVL-encoding phages—ΦPVL or ΦPVL-like (*n* = 6, 19.4%), Φ108PVL or Φ108PVL-like (*n* = 15, 48.4%), ΦTCH60 (*n* = 1, 3.2%), and ΦSa2USA (*n* = 1, 3.2%), whereas the remainder (*n* = 8, 25.8%) harbored an unidentifiable phage with icosahedral head morphology (Table 1). In contrast, SWP strains carried ΦTCH60 (*n* = 10, 71.4%), ΦSa2958 (*n* = 3, 21.5%), and Φ108PVL (*n* = 1, 7.1%), while Con strains harbored either ΦSa2USA (*n* = 4, 57.1%), ΦSa2USA/ΦSa2958 (*n* = 2, 28.6%), or ΦPVL (*n* = 1, 14.3%).

PVL gene variants. Using the PVL gene of ΦPVL108 as a reference sequence, four SNPs were identified among the 52 *lukS-PV* sequences, while five were identified in *lukF-PV* (Table 1). Of the nine SNPs, four were nonsynonymous, including a previously described arginine-to-histidine replacement at amino acid residue 176 (nucleotide [nt] 527) (*n* = 6) (14) and a phenylalanine-to-tyrosine replacement at amino acid 157 (nt 470) (*n* = 5) (13). An additional two nonsynonymous mutations were identified, including a glycine-to-alanine replacement at amino acid 99 (nt 296) in *lukS-PV* (*n* = 1) and an alanine-to-valine replacement at amino acid 139 (nt 1356) in *lukF-PV* (*n* = 1).

O'Hara et al. (14) assigned different *lukSF-PV* alleles into four major haplotype groups (H1, H2, H3, and R) on the basis of variation at nucleotide positions 527, 663, and 1396; the same nomenclature has been used in the present study. To explore the nucleotide sequence differences within each haplotype group, we further assigned each unique *lukSF-PV* gene variant to different haplotype alleles, i.e., H1a, H1b, H1c, R1, R2, R3, etc. (Table 1 and Fig. 1). In the present study, a total of 10 unique haplotype alleles were identified, with haplotype H2a (*n* = 26, 50.0%) and H1a (*n* = 9, 17.3%) representing the most common types.

We further examined the relationships between PVL haplotype alleles, PVL-encoding phages, and CC30 genetic lineage. In the present study, the H2b allele was mainly associated with Φ108PVL (16/26, 57.7%) and ΦTCH60 (9/26, 34.6%), while H1a was mostly observed within ΦPVL (5/9, 55.6%) and ΦSa2958 (3/9, 33.3%).

FIG 1 Statistical parsimony network constructed using TCS 1.21 (28), depicting the phylogenetic relationship of *lukSF-PV* genes among 52 *S. aureus* CC30 isolates, as well as CC8 strain USA300-FPR3757. USA300-FPR3757 is used as a reference strain, and the dotted line indicates the relationship of *lukSF-PV* genes between USA300-FPR3757 and the CC30 isolates. Ellipses indicate discrete *lukSF-PV* alleles; ellipse size is proportional to the frequency of the allele (except for USA300-FPR3757). The rectangular box indicates the predicted ancestral allele. Lines connecting ellipses indicate single nucleotide substitutions. The small white circles between H1a and R2 and between H2a and R1 denote intermediate haplotypes which were not observed in this collection. R haplotype variants are shaded in gray. PVL phages associated with specific haplotypes are shown alongside the ellipses.

The H1b allele, by contrast, was exclusively found in strains harboring a nontypeable icosahedral head phage, isolated from New Jersey and New York in 2004 and 2007. Similarly, the H2b allele was associated with nontypeable icosahedral head phage-harboring strains isolated in France and the United States from 2002 to 2007. Interestingly, all six R haplotype (R1, R2, and R3)-harboring strains were positive for the Φ Sa2USA-specific PCR, while two of them were also positive for the Φ Sa2958-specific PCR (Table 1).

To reconstruct the evolutionary history of the PVL genes, we used TCS software to estimate a phylogenetic network of 52 PVL sequences using the method described previously (14) (Fig. 1). On the basis of statistical parsimony, the H variant is predicted to be the ancestor of the R variant, in agreement with previous results from an international collection of 174 clinical *S. aureus* strains with different genetic backgrounds, including CC1, CC5, CC8, CC22, CC30, and CC121, etc. (14). The H1a allele can change into the R2 type via two single nucleotide substitutions or, alternatively, the H2a allele can mutate through an intermediate H3 allele into the R1 form, which is identical to the *lukSF-PV* sequence of MW2. The *lukSF-PV* sequence of USA300-FPR3757, designated haplotype R0, contains a single nucleotide difference relative to the R1 type (Fig. 1).

DNA sequencing of phage/chromosome junctions. In the present study, all 52 isolates showed positive amplification for a PCR assay designed to detect the proximal and distal junctions of PVL phage insertion sites in CC30 strains (16), suggesting a common PVL phage integration site regardless of difference among CC30 sublineages, PVL phage types, and PVL gene variants. DNA

sequencing of the junction regions revealed that all strains share the same left attachment site (*attL*), showing 100% identity to *attL* in strains 81/108 (Φ PVL108) and TCH60 (Φ TCH60). In contrast, the sequences of the right attachment sites (*attR*) were slightly more diverse, displaying unique mutations at four positions. Maximum-likelihood tree analysis of *attR* sequences divided them into two clusters (A and B) and two singletons (C and D). Within the two clusters, 41 strains were assigned into cluster A, whereas 9 strains were assigned into cluster B (Table 1). The *attR* sequences of cluster A strains were identical to those of strain 81/108, whereas the *attR* sequences of cluster B were the same as those of strain TCH60. Notably, the 9 cluster B strains all possessed the same *spa* type (eGenomics/Ridom *spa* type 19/t019), harbored the same PVL phage (Φ TCH60), and displayed similar PVL haplotypes (H2a and H2c). With the exception of one strain (BK21010), the remaining 8 strains harbored an SCCmec IV element and were therefore classified as belonging to the SWP clone.

DISCUSSION

PVL-positive *S. aureus* has acquired significant attention in recent years, primarily due to its epidemiologic association with CA-MRSA, including the widespread USA300 clone (31). In the present study, the relationships between CC30 clones, PVL phages, PVL variant alleles, and the chromosomal phage insertion junctions were explored using a collection of 52 geographically and temporally diverse CC30 strains in order to better understand the evolution of PVL within this pandemic lineage.

Both icosahedral-head and elongated-head phages were observed in all three sublineages of CC30; however, the pandemic phage 80/81 strains were mainly associated with icosahedral-head phages, whereas the SWP and Con strains were more likely to harbor elongated phages (Table 1). This finding is in agreement with previous comparative genome studies, suggesting that the progenitor CC30 clones initially harbored an icosahedral-head phage encoding PVL (2). Presumably, the icosahedral-head phage was subsequently replaced by the elongated-head phages present in most isolates of the SWP and Con clones. The presence of multiple PVL phages is suggestive of multiple independent acquisition or horizontal transmission of PVL phages within CC30 sublineages. Our results also demonstrate that PVL has displayed a long residency within some community-associated *S. aureus* clones (e.g., phage 80/81) and has been maintained within the CC30 lineage for at least 77 years (Table 1). These data also underline the hypothesis that PVL-harboring CC30 strains may constitute an important reservoir for the transmission of PVL phages.

The highly conserved sequences observed within the *lukSF-PV* genes themselves indicate that they all share a recent common ancestor (14). Within the Con CC30 clone, we identified six R haplotype PVL gene variants which have been primarily associated with CA-MRSA strains, especially the USA300 CC8 clone (14). Interestingly, all R haplotype-associated strains carry a phage resembling the Φ Sa2USA phage described in USA300 strains emerging in the early 2000s (Table 1). Two R-type strains also carry an element similar to Φ Sa2958, suggesting the presence of a mosaic phage structure, or the coexistence of both Φ Sa2958-like and Φ Sa2USA-like phages within a single strain. In a previous study, O'Hara et al. reconstructed the origins of the PVL gene in CA-MRSA strain USA300 (R0 type PVL in Fig. 1), involving transmission of PVL from CA-MRSA strain MW2 (USA400) to a CC8 SCCmec IV-harboring strain (14). Interestingly, our data demon-

strate that the R-type *lukSF-PV* haplotype existed as early as 1986, prior to the emergence of PVL-harboring CA-MRSA strains (32–35). The association between the R haplotype and phage Φ Sa2USA within CC30 strains further suggests that the PVL genes in USA300 may be derived from a CC8 strain that acquired Φ Sa2USA from a CC30 Con strain, either before or after the acquisition of SCCmec IV. Further studies, however, including sequencing of the Con R phages, are needed to confirm this hypothesis.

Another interesting finding is that all CC30 strains share the same insertion site for the various PVL phages identified here, thereby exhibiting the same CC30 lineage-specific site described in a previous study (16). In the latter study, Boakes et al. demonstrated the presence of two lineage-specific insertion sites within the *S. aureus* chromosome: Ins1 in CC22, -30, and -80 and Ins2 in CC1, -5, -8, -59, -88, and -93 and occasionally CC22 (16). Moreover, site-specific attachment sequences (*attL* and *attR*) are generally conserved across different lineages of PVL-harboring MRSA strains (16). Our study further confirms that distinct CC30 clones (80/81, SWP, and Con) possess the same insertion site, in agreement with the findings of Boakes et al. (16). The main variations in *attR* are located in the 29-bp chromosome binding sites (B2), indicating that the diversity of *attR* sequences are due to differences in the chromosomal DNA. The strains in cluster B possess highly similar genetic backgrounds, PVL variants and PVL phage head types, suggestive of a recent evolutionary branching of SWP strains. Based on two SNPs (*hla*_{STOP} and *agr*C_{G55R}) and the prevalence of SCCmec IV, strain 21010 has been tentatively assigned to the 80/81 lineage. However, based on the *attR* sequence, as well as a similar genotypic background compared to other members of cluster B (e.g., *spa* type, PVL phage, and PVL gene haplotype) (Table 1), it is likely that 21010 is a methicillin-susceptible SWP strain. Additional sequence analysis of the SCCmec region will be required to understand whether this strain is a “true” MSSA or an MRSA that subsequently lost the SCCmec IV element.

This is the first study detailing the population structure of PVL-harboring CC30 clones, which expands our current knowledge of the evolutionary relationships between distinct CC30 sublineages. Both PVL gene sequences and PVL-encoding phages vary within sublineages of PVL-positive CC30 strains, while the observation of the PVL R-haplotype in CC30 strains sheds further light on the evolution of PVL-positive CA-MRSA.

REFERENCES

- Robinson DA, Kearns AM, Holmes A, Morrison D, Grundmann H, Edwards G, O'Brien FG, Tenover FC, McDougal LK, Monk AB, Enright MC. 2005. Re-emergence of early pandemic *Staphylococcus aureus* as a community-acquired methicillin-resistant clone. *Lancet* 365:1256–1258.
- McAdam PR, Templeton KE, Edwards GF, Holden MT, Feil EJ, Aanensen DM, Bargawi HJ, Spratt BG, Bentley SD, Parkhill J, Enright MC, Holmes A, Girvan EK, Godfrey PA, Feldgarden M, Kearns AM, Rambaut A, Robinson DA, Fitzgerald JR. 2012. Molecular tracing of the emergence, adaptation, and transmission of hospital-associated methicillin-resistant *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U. S. A.* 109: 9107–9112.
- Chambers HF, Deleo FR. 2009. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat. Rev. Microbiol.* 7:629–641.
- DeLeo FR, Kennedy AD, Chen L, Bubeck Wardenburg J, Kobayashi SD, Mathema B, Braughton KR, Whitney AR, Villaruz AE, Martens CA, Porcella SF, McGavin MJ, Otto M, Musser JM, Kreiswirth BN. 2011. Molecular differentiation of historic phage-type 80/81 and contemporary epidemic *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U. S. A.* 108:18091–18096.
- Panton PN, Valentine FCO. 1932. Staphylococcal toxin. *Lancet* 219:506–508.
- Kaneko J, Kimura T, Narita S, Tomita T, Kamio Y. 1998. Complete nucleotide sequence and molecular characterization of the temperate staphylococcal bacteriophage Φ PVL carrying Panton-Valentine leukocidin genes. *Gene* 215:57–67.
- Ma XX, Ito T, Chongtrakool P, Hiramatsu K. 2006. Predominance of clones carrying Panton-Valentine leukocidin genes among methicillin-resistant *Staphylococcus aureus* strains isolated in Japanese hospitals from 1979 to 1985. *J. Clin. Microbiol.* 44:4515–4527.
- Zhang M, Ito T, Li S, Jin J, Takeuchi F, Lauderdale TL, Higashide M, Hiramatsu K. 2011. Identification of the third type of PVL phage in ST59 methicillin-resistant *Staphylococcus aureus* (MRSA) strains. *FEMS Microbiol. Lett.* 323:20–28.
- Ma XX, Ito T, Kondo Y, Cho M, Yoshizawa Y, Kaneko J, Katai A, Higashiide M, Li S, Hiramatsu K. 2008. Two different Panton-Valentine leukocidin phage lineages predominate in Japan. *J. Clin. Microbiol.* 46: 3246–3258.
- Baba T, Takeuchi F, Kuroda M, Yuzawa H, Aoki K, Oguchi A, Nagai Y, Iwama N, Asano K, Naimi T, Kuroda H, Cui L, Yamamoto K, Hiramatsu K. 2002. Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* 359:1819–1827.
- Narita S, Kaneko J, Chiba J, Piemont Y, Jarraud S, Etienne J, Kamio Y. 2001. Phage conversion of Panton-Valentine leukocidin in *Staphylococcus aureus*: molecular analysis of a PVL-converting phage, Φ SLT. *Gene* 268: 195–206.
- Diep BA, Gill SR, Chang RF, Phan TH, Chen JH, Davidson MG, Lin F, Lin J, Carleton HA, Mongodin EF, Sensabaugh GF, Perdreau-Remington F. 2006. Complete genome sequence of USA300, an epidemic clone of community-acquired methicillin-resistant *Staphylococcus aureus*. *Lancet* 367:731–739.
- Takano T, Higuchi W, Otsuka T, Baranovich T, Enany S, Saito K, Isobe H, Dohmae S, Ozaki K, Takano M, Iwao Y, Shibuya M, Okubo T, Yabe S, Shi D, Reva I, Teng LJ, Yamamoto T. 2008. Novel characteristics of community-acquired methicillin-resistant *Staphylococcus aureus* strains belonging to multilocus sequence type 59 in Taiwan. *Antimicrob. Agents Chemother.* 52:837–845.
- O'Hara FP, Guex N, Word JM, Miller LA, Becker JA, Walsh SL, Scangarella NE, West JM, Shawar RM, Amrine-Madsen H. 2008. A geographic variant of the *Staphylococcus aureus* Panton-Valentine leukocidin toxin and the origin of community-associated methicillin-resistant *S. aureus* USA300. *J. Infect. Dis.* 197:187–194.
- Li X, Sun J, Wu D, Wang L, Yang Y, Wang C, Liu L, Zhao C, Du P, Yu S, Shen X. 2012. Panton-Valentine leukocidin gene sequence variation and phage in methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* from children in mainland China. *Microbiol. Immunol.* 56:155–162.
- Boakes E, Kearns AM, Ganner M, Perry C, Hill RL, Ellington MJ. 2011. Distinct bacteriophages encoding Panton-Valentine leukocidin (PVL) among international methicillin-resistant *Staphylococcus aureus* clones harboring PVL. *J. Clin. Microbiol.* 49:684–692.
- Brown ML, O'Hara FP, Close NM, Mera RM, Miller LA, Suaya JA, Amrine-Madsen H. 2012. Prevalence and sequence variation of Panton-Valentine leukocidin in methicillin-resistant and methicillin-susceptible *staphylococcus aureus* strains in the United States. *J. Clin. Microbiol.* 50: 86–90.
- Otter JA, Kearns AM, French GL, Ellington MJ. 2010. Panton-Valentine leukocidin-encoding bacteriophage and gene sequence variation in community-associated methicillin-resistant *Staphylococcus aureus*. *Clin. Microbiol. Infect.* 16:68–73.
- Wolter DJ, Tenover FC, Goering RV. 2007. Allelic variation in genes encoding Panton-Valentine leukocidin from community-associated *Staphylococcus aureus*. *Clin. Microbiol. Infect.* 13:827–830.
- Coombs GW, Goering RV, Chua KY, Monecke S, Howden BP, Stinear TP, Ehrlich R, O'Brien FG, Christiansen KJ. 2012. The molecular epidemiology of the highly virulent ST93 Australian community *Staphylococcus aureus* strain. *PLoS One* 7:e43037. doi:10.1371/journal.pone.0043037.
- Boakes E, Kearns AM, Ganner M, Perry C, Warner M, Hill RL, Ellington MJ. 2011. Molecular diversity within clonal complex 22 methicillin-resistant *Staphylococcus aureus* encoding Panton-Valentine leukocidin in England and Wales. *Clin. Microbiol. Infect.* 17:140–145.
- Chen L, Mediavilla JR, Smyth DS, Chavda KD, Ionescu R, Roberts RB, Robinson DA, Kreiswirth BN. 2010. Identification of a novel transposon

- (Tn6072) and a truncated SCCmec element in methicillin-resistant *Staphylococcus aureus* ST239. Antimicrob. Agents Chemother. 54:3347–3354.
23. Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG. 2000. Multi-locus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. J. Clin. Microbiol. 38:1008–1015.
  24. Mathema B, Mediavilla J, Kreiswirth BN. 2008. Sequence analysis of the variable number tandem repeat in *Staphylococcus aureus* protein A gene: spa typing. Methods Mol. Biol. 431:285–305.
  25. Chen L, Mediavilla JR, Oliveira DC, Willey BM, de Lencastre H, Kreiswirth BN. 2009. Multiplex real-time PCR for rapid staphylococcal cassette chromosome mec (SCCmec) typing. J. Clin. Microbiol. 47:3692–3706.
  26. Milheirico C, Oliveira DC, de Lencastre H. 2007. Multiplex PCR strategy for subtyping the staphylococcal cassette chromosome mec type IV in methicillin-resistant *Staphylococcus aureus*: “SCCmec IV multiplex.” J. Antimicrob. Chemother. 60:42–48.
  27. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. CLUSTAL W and CLUSTAL X version 2.0. Bioinformatics 23:2947–2948.
  28. Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9:1657–1659.
  29. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum-parsimony methods. Mol. Biol. Evol. 28:2731–2739.
  30. Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22:160–174.
  31. Mediavilla JR, Chen L, Mathema B, Kreiswirth BN. 2012. Global epidemiology of community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA). Curr. Opin. Microbiol. 15:588–595.
  32. Herold BC, Immergluck LC, Maranan MC, Lauderdale DS, Gaskin RE, Boyle-Vavra S, Leitch CD, Daum RS. 1998. Community-acquired methicillin-resistant *Staphylococcus aureus* in children with no identified predisposing risk. JAMA 279:593–598.
  33. Dufour P, Gillet Y, Bes M, Lina G, Vandenesch F, Floret D, Etienne J, Richet H. 2002. Community-acquired methicillin-resistant *Staphylococcus aureus* infections in France: emergence of a single clone that produces Pantone-Valentine leukocidin. Clin. Infect. Dis. 35:819–824.
  34. Vandenesch F, Naimi T, Enright MC, Lina G, Nimmo GR, Heffernan H, Liassine N, Bes M, Greenland T, Reverdy ME, Etienne J. 2003. Community-acquired methicillin-resistant *Staphylococcus aureus* carrying Pantone-Valentine leukocidin genes: worldwide emergence. Emerg. Infect. Dis. 9:978–984.
  35. Diep BA, Sensabaugh GF, Somboonna N, Carleton HA, Perdreau-Remington F. 2004. Widespread skin and soft-tissue infections due to two methicillin-resistant *Staphylococcus aureus* strains harboring the genes for Pantone-Valentine leukocidin. J. Clin. Microbiol. 42:2080–2084.