JOURNAL OF BACTERIOLOGY, Jan. 2006, p. 169–178 Vol. 188, No. 1
0021-9193/06/$08.00+0 doi:10.1128/JB.188.1.169–178.2006

Evidence for Clonal Evolution among Highly Polymorphic Genes in Methicillin-Resistant *Staphylococcus aureus*

G. Kuhn,* P. Francioli, and D. S. Blanc

Hospital Preventive Medicine, University Hospital of Lausanne (CHUV), Lausanne, Switzerland

Received 29 July 2005/Accepted 7 October 2005

The evolution of *Staphylococcus aureus* has been described as predominantly clonal, based on evidence from seven housekeeping genes. We aimed to test if this was also true for more polymorphic genes. In a collection of 60 isolates including major European epidemic methicillin-resistant *S. aureus* (MRSA) and sporadic MRSA strains, we compared the partial gene sequences of seven housekeeping genes (*arcC*, *aroE*, *glpF*, *gmk*, *pta*, *tpi*, and *yqiL*), six core adhesion genes (present in all strains) (*clfA*, *clfB*, *fnbA*, *map*, *sdrC*, and *spa*), and four accessory adhesion genes (not present in all strains) (*ebpS*, *fnbB*, *sdrD*, and *sdrE*). Nucleotide diversity of adhesion genes was 2- to 10-fold higher than genes used for multilocus sequence typing. All genes showed evidence for purifying selection with a weakly reduced level among accessory adhesion genes. Among these highly variable genes, there was no evidence for a difference in molecular evolution between epidemic and sporadic strains. Gene trees constructed from concatenated sequences of housekeeping, core adhesion, and accessory adhesion genes were highly congruent, indicating clonality, despite some evidence for homologous exchange. Further evidence for clonality was found with an overall positive correlation of allelic and nucleotidic divergence for both seven housekeeping genes and six core adhesion genes. However, for small allelic differences that fit the demarcations of clonal complexes (CCs) there was no such correlation, suggesting that recombination occurred. Therefore, despite an overall clonal population structure, recombination between related isolates within CCs might have contributed to *S. aureus* evolution.

Methicillin-resistant *Staphylococcus aureus* (MRSA) is a major cause of nosocomial infections worldwide. The health conditions caused by MRSA can range from carriage without symptoms to severe life-threatening infections. Epidemiological studies have revealed that a high number of these infections were caused by single MRSA clones that spread rapidly in hospital settings (epidemic MRSA), while other clones that were present at the same time did not spread (sporadic MRSA) (6, 7, 8, 30). Genetic variation causes phenotypic differences, and population genetic approaches can be used to identify how this variation is created and maintained.

Multilocus sequence typing (MLST) has been developed to analyze and compare genetic variation in worldwide collections of bacterial isolates (27). Using the MLST sequence data of *S. aureus*, it was shown that genetic differences between a single-locus variant and its ancestral strains were created 15 times more frequently by a mutation than by a recombination event (14). Clearly, this indicated a predominantly clonal evolution.

MLST has been based on seven slowly evolving genes that encode proteins of central metabolic functions (housekeeping) (11), and it was used to analyze the evolutionary history of MRSA (12). A more polymorphic set of genes, the *S. aureus* cell surface genes (*sas*) that are mostly scattered around the origin of replication, have been analyzed as an extension of this typing scheme and resolved a more-detailed picture of MRSA evolution (32). Both sets showed a high congruence in their phylogenetic resolution of groups, which again suggested clonal evolution.

* Corresponding author. Mailing address: Hospital Preventive Medicine, University Hospital of Lausanne (CHUV), Avenue du Bugnon 48, 1011 Lausanne, Switzerland. Phone: 41 21 314 0268. Fax: 41 21 314 4060. E-mail: Gerrit.Kuhn@chuv.ch.

Recently, bacterial genomes were suggested to be viewed as two compartments of genes, one comprising "core" genes that are ubiquitously present in all clones of a given species and the other comprising "accessory" genes that are not necessarily present in all clones and that have the possibility to be transferred between strains (24). MLST is based on seven housekeeping genes that represent exclusively core genome genes; because of their definition, these genes can only evolve through mutation and gene replacement. Accessory genes also have potential to get lost and to be acquired during evolution (they are therefore considered "dispensable"), and thus the resulting population structure given by these genes might be different from that of core genome genes (13).

Comparative genome hybridization experiments were performed to address genetic variation due to the presence and absence of genes in a collection of highly diverse *S. aureus* strains. This study showed that up to 22% of all genes in the genome could be considered as accessory (16). Thus, there is a considerable proportion of the genome that seemed not to be essential and that had the potential to contribute to genetic differences between strains.

The study of accessory genes is of utmost importance since they include resistance to methicillin (*mecA*) (12) and virulence factors (*sea*, *tst*, and *eta*) (3, 26, 28, 41). To date, only the molecular evolution of locus (RD13), comprising an exotoxin-like family of genes, was analyzed. This locus likely evolved clonally through multiple gene deletions, and genes were subject to purifying selection (15).

The goal of the present study was to compare the evolution of "conserved" genes and polymorphic genes among MRSA isolates. We selected the seven housekeeping genes used for MLST analysis as representatives of conserved genes (11). For

TABLE 1. Details of typing analyses of 60 *S. aureus* isolates from different European countries, grouped according to MLST housekeeping typing data (CC and ST)

Study no.	Original no.	Yr or date (mo. day. yr) of isolation	Origin	Clone name	Epidemiology[a]	CC[b]	ST[c]	Ridom[d]	RAPD[e]	Gene										Reference(s) or source
										clfA	*clfB*	*ebpS*	*fnbA*	*fnbB*	*map*	*sdrC*	*sdrD*	*sdrE*	*spa*	
MSSA476	MSSA476	1998	England			1	1	New		5	4	4	13	7	5	9	13	10	15	19
MW2	MW2	1998	United States			1	1	t128		5	4	4	13	7	4	10	12	10	15	2
45	H9779	2000	CHUV[f]		Sporadic	1	New	t127		5	4	4	13	7	3	9	11	10	14	This study
41	H6556	1996	CHUV[f]		Sporadic	5	5	t045		1	1	2	17	1	1	1	1	2	7	This study
43	H7920	1998	CHUV[f]		Sporadic	5	5	New		1	1	1	2	NA[h]	1	1	1	1	4	This study
80	3810/5	1995	Belgium		Sporadic	5	5	New		1	1	1	2	1	1	1	1	2	1	6
Mu50	Mu50	1996	Japan			5	5	t002		1	1	2	2	1	1	1	1	2	7	23
N315	N315	1982	Japan			5	5	t002		1	1	2	2	1	1	1	1	2	7	23
47	MR3	1992	Poland	Clone A4		5	83	t053	1	1	1	1	19	1	1	1	1	3	3	4
1	H9140	10.06.1999	CHUV[f], pt1	B	Epidemic	5	105	t002		1	1	1	1	1	1	1	1	1	1	This study
2	H13199	05.09.2003	CHUV[f], pt1	B		5	105	t002		1	1	1	1	1	1	1	1	1	5	This study
34	IH99802	2001	Finland	Bel EC-3	Epidemic	5	125	t067		1	1	1	18	1	1	1	1	1	2	
3	H9502	11.01.2000	CHUV[f], pt2	D	Epidemic	5	228	t041		1	1	1	2	NA	1	1	2	1	6	This study
4	H13442	01.12.2003	CHUV[f], pt2	D		5	228	t041		1	1	1	2	NA	1	1	2	1	6	This study
81	7102/1	1995	Belgium		Sporadic	5	228	t001		1	1	1	2	NA	1	1	1	2	6	
70	ZH224	1997	Switzerland			8	8	t064	2b	11	1	13	3	10	13	14	15	14	20	4
82	no 15 Mergan	1999	Belgium		Sporadic	8	8	t207		11	1	13	3	10	13	14	NA	NA	20	
83	A930016	1993	France	Gm-S[g]	Epidemic	8	8	t207		11	1	13	3	10	13	14	NA	NA	20	5
84	A930018	1993	France	Gm-S[g]	Epidemic	8	8	t207		11	1	13	3	10	13	14	NA	NA	20	5
85	A930021	1993	France	Gm-S[g]	Epidemic	8	8	t207		11	1	13	3	10	13	14	NA	NA	20	5
93	H11398	2001	CHUV[f]		Sporadic	8	8	t207		11	1	13	3	10	13	14	NA	NA	20	This study
NCTC8325	NCTC8325	1963	United States			8	8	t211		13	1	13	6	10	13	14	15	NA	20	20
26	HSJ216	1997	Portugal	Brazilian	Epidemic	8	239	t037		10	6	13	3	10	13	13	15	13	8	8
27	HSJ7	1995	Portugal	Brazilian	Epidemic	8	239	t037		10	6	13	3	10	13	13	15	13	8	8
28	HUC77	1994	Portugal	Brazilian	Epidemic	8	239	t037		10	6	13	3	10	13	13	15	13	8	29
42	H7051	1996	CHUV[f]		Sporadic	8	239	t030		10	6	13	4	NA	11	14	15	14	10	6
46	H7951	1998	CHUV[f]		Sporadic	8	239	t030		10	6	13	3	NA	11	14	15	14	10	This study
68	H7681	1997	Switzerland			8	239	t037	2a	10	6	13	5	10	12	14	15	14	8	4
72	GE94	1997	Switzerland			8	239	t037	2a	10	5	13	3	NA	13	14	15	14	8	4
55	635/93	1993	Austria	Phage group III		8	New	t037	2a	10	6	13	3	10	13	13	15	13	8	4
23	HPV107	1992	Portugal	Iberian	Epidemic	8	247	t051		9	1	12	3	10	13	14	14	NA	20	35
24	VNG2	1992	Portugal	Iberian	Epidemic	8	247	t051		9	1	12	3	10	13	14	14	NA	20	34
25	HUC9	1992	Portugal	Iberian	Epidemic	8	247	t051		9	1	12	3	10	13	14	14	NA	20	29
58	82	1993	Belgium	Clone 1		8	247	t054	2b	9	1	12	3	10	13	14	14	12	20	4
78	2104/3	1995	Belgium		Sporadic	8	247	New		9	1	12	3	10	13	14	14	14	20	
79	5301/5	1995	Belgium		Sporadic	8	247	t052		9	1	12	3	10	13	14	14	11	20	
COL	COL	1960's	England			8	250	t008		11	1	13	3	10	13	14	15	14	20	17
54	994/93	1993	Germany	Phage group III		8	254	t139	2b	12	1	13	7	10	13	14	15	NA	20	4
33	IH98929	1999	Finland	Kokkola	Epidemic	8	New	New		11	1	14	3	10	13	14	15	NA	21	
29	IPOP2	2001	Portugal	EMRSA-15	Epidemic	22	22	t032		3	3	11	13	11	7	2	4	9	16	9
64	4C6F		England	EMRSA-15		22	22	New	3	3	3	11	13	NA	7	2	4	9	17	4
30	HGSA128	2000	Portugal	EMRSA-15	Epidemic	22	79	t032		3	3	11	13	12	7	2	4	9	16	9
66	H7531	1997	Switzerland			30	30	t268	4	8	6	10	10	NA	10	7	6	6	9	4
74	GE149	1997	Switzerland			30	30	t037	4	8	6	9	10	2	10	8	5	7	8	4
63	4C7E	1993	England	EMRSA-16		30	36	t268	4	8	6	9	10	NA	10	7	5	6	9	4
MRSA252	MRSA252	1997	England	EMRSA-16		30	36	t018		8	6	9	10	NA	10	7	NA	6	9	19
31	H6606	1996	CHUV[f]	A	Epidemic	45	45	t038		6	8	5	14	4	8	3	8	8	11	4, 6
32	IH76167	1996	Finland	Kemi	Epidemic	45	45	t015		6	8	5	14	5	8	3	8	8	11	
38	01-1111	2001	Netherlands		Epidemic	45	45	t445		6	8	5	14	5	8	3	8	8	11	40

39	01-1222	2001	Netherlands		Epidemic	45	45	t445		6	8	5	14	5	8	3	8	8	11	40	
40	01-1279	2001	Netherlands		Epidemic	45	45	t445		6	8	5	14	5	8	3	8	8	11	40	
56	1150/93	1993	Germany	Phage pattern 95		45	45	t004	5	6	7	6	16	4	8	5	9	8	12	4	
71	GE79	1997	Switzerland	GE-3		45	45	New	6	7	8	8	11	8	9	6	10	8	13	4	
92	H13717	2004	CHUV[f]		Sporadic	45	45	New		7	8	8	12	9	9	6	NA	NA	13	4	This study
36	B1	2003	Norway		Epidemic	45	455	t065		6	9	5	15	5	8	4	8	8	12	25	
37	B2	2003	Norway		Epidemic	45	455	t065		6	9	5	15	5	8	4	8	8	12	25	
35	A1	2003	Norway		Epidemic	45	497	New		6	8	5	14	5	8	3	8	8	12	25	
62	45030564	1996	Canada	OE-MRSA		45	New	t004	5	6	9	7	14	4	8	3	8	8	12	4	
90	H13911	2004	CHUV[f]		Sporadic	S	15	t279		2	4	NA	8	3	6	12	3	5	19	This study	
44	H7639	1997	CHUV[f]		Sporadic	S	80	t131		4	2	3	9	6	2	11	7	4	18	6	

[a] Isolates were called epidemic if they were sampled at the beginning of an outbreak. They were considered sporadic if they were never found to be transmitted to other patients in a defined epidemiological setting.
[b] Sequence types (STs) were assigned to CCs, according to previously published data (14, 32). S, singletons.
[c] STs were determined with the MLST database located at http://saureus.mlst.net/.
[d] Ridom *spa* types were determined with the Ridom *spa* server located at http://www.ridom.de/*spa*server/.
[e] RAPD groups were determined in a previous analysis by a multiprimer RAPD approach (4).
[f] University Hospital of Lausanne.
[g] Gentamycin-susceptible clone.
[h] NA, no amplification with PCR.

polymorphic genes, we selected cell surface genes; within these, we aimed to differentiate core cell surface genes and accessory cell surface genes. We selected cell surface genes because in contrast to the slowly evolving genomic core genes this class of genes has been shown to undergo rapid evolutionary changes, which already rendered them useful for short-term epidemiological typing (18, 21, 36). In addition, adhesion to host tissues and factors therein is important for successful host colonization; thus, these genes allow us to test the hypothesis of whether a specific genetic signature or molecular evolutionary pattern is linked to epidemic potential.

By sequencing the DNA of 17 different genes from a collection of 60 isolates, we aimed to study *S. aureus* genome evolution. In particular, we aimed to study molecular evolutionary patterns of 10 adhesion genes (*clfA*, *clfB*, *ebpS*, *fnbA*, *fnbB*, *map*, *sdrC*, *sdrD*, *sdrE*, and *spa*) and compare these to seven housekeeping genes (*arcC*, *aroE*, *glpF*, *gmk*, *pta*, *tpi*, and *yqiL*).

MATERIALS AND METHODS

Bacterial isolates. A total of 60 isolates was used in this study (Table 1). Fourteen of these isolates were chosen because they represent members of the six MRSA clusters present in Europe that were previously described (4). The other strains were selected to study the evolution of epidemicity in MRSA. Our study aimed to confront two different groups of isolates, those that were shown to disseminate rapidly in a setting and those that were never reported to disseminate in a setting. Twenty-three MRSA isolates belonged to 12 different epidemic clones. Each of these isolates was collected at the onset of an epidemic, because we aimed to exclude isolates that might have undergone genetic changes during the epidemic, which could have affected their epidemic potential. For six of these clones (EMRSA-15, Norway, Iberian, Brazilian, The Netherlands, and Gm-S) (Table 1), we had either two or three isolates. They had the same genotype (see the references cited in Table 1) but were sampled at different locations. For two clones (B and D), we had two additional isolates that were obtained from the same patient after a 4-year period of time (isolates 2 and 4). Previous genotyping analysis using pulsed-field gel electrophoresis showed that the first and the second isolates from each patient were identical. Fourteen isolates were considered sporadic, because they were never found to be transmitted between patients, because they were never isolated from another patient in the same setting, and because they differed in pulsed-field gel electrophoresis patterns from epidemic clones. In addition, sporadic CHUV isolates were present on the patients for at least 3 weeks during hospitalization and thus had the possibility of disseminating. The sequences from the seven publicly available genome sequences were also included.

Cultivation of bacteria, lysis, and DNA extraction. *S. aureus* isolates were grown in 1 ml brain heart infusion liquid medium. Pelleted cells were lysed in 200 μl lysis buffer (1× Tris-EDTA buffer, 0.35 M NaCl, and 0.05-mg/ml lysostaphin). DNA was extracted from 100 μl lysate as described previously (10).

Design of 10 adhesion gene primers. Gene sequences of published genes *clfA* (Z18852), *clfB* (AJ224764), *ebpS* (U48826), *fnbA* (J04151), *fnbB* (X62992), *map* (AJ245439), *sdrC* (AJ005645), *sdrD* (AJ005646), *sdrE* (AJ005647), and *spa* (M18264) were taken for a BLAST analysis with the seven publicly available *S. aureus* genome sequences to obtain their homologous gene sequences (COL, NCTC8325, MW2, methicillin-susceptible *Staphylococcus aureus* 476 [MSSA476], MRSA252, Mu50, and N315). Sequences of each gene were aligned, and primers were targeted against regions with no genetic variation.

The genes *sdrC*, *sdrD*, and *sdrE* shared regions with high sequence similarity among each other, leading to cross-hybridization in first PCR trials. The primers were redesigned to match conserved regions that were specific for each of the three genes. Table 2 summarizes the names, the supposed or known function of the gene, the size of the amplicons, and both primer names and sequences.

PCR and sequencing. PCR amplifications for each gene were carried out in 25-μl reaction volumes containing 40 ng of bacterial DNA, 0.5 μM of each primer, 1 U of *Taq* DNA polymerase (Invitrogen), 1× PCR buffer, 1.5 mM $MgCl_2$ (supplied with the *Taq* polymerase), and 0.1 mM deoxynucleotide triphosphates (Invitrogen). Cycling conditions were 3 min of initial denaturation at 94°C; 32 cycles consisting of 30 s at 94°C, and 30 s at primer-annealing temperature (Table 2), 1 min (2 min, if amplicons were <1,000 bp) at 72°C; and a final extension for 5 min at 72°C.

Standard gel electrophoresis was performed, and stained gels were evaluated

TABLE 2. Details of the 10 adhesion genes and the regions amplified and sequenced in this study

Gene	Supposed or known function	Accession no. of reference sequence	Primer name	Sequence (5′–3′)	Annealing temp (°C)[a]	Position on reference sequence
clfA	Clumping factor A; fibrinogen receptor	Z18852	SA_clfA_671.for	AAGCTAATACACCGGCAACAACTC	56	671
			SA_clfA_1943.rev	GTTCAATTTCACCAGGCTCATCAG	56	1943
clfB	Clumping factor B; fibrinogen receptor	AJ224764	SA_clfB_353.for	AACAGAGCCAGCTTCAACAAATGA	56	353
			SA_clfB_1638.rev	GTCTTTCGGATTTACTGCTGAATC	56	1638
ebpS	Elastin-binding protein	U48826	SA_ebpS_61.for	ATCAAGACCATACGGAAGATGTTG	56	61
			SA_ebpS_1216.rev	GTGCTTGATTGTTCGCTTGATTAG	56	1216
fnbA	Fibronectin-binding protein A	J04151	SA_fnbA_420.for	ACCACAAACTGCACAACCAGCA	50	420
			SA_fnbA_1868.rev	CTATCCACAGCAGTATGGTAATCG	50	1868
fnbB	Fibronectin-binding protein B	X62992	SA_fnbB_233.for	CGACATCAACTGAGCAACCATC	50	233
			SA_fnbB_1565.rev	TCTGCACCTTCAACAGCTGTATGA	50	1565
map	MHC class II analog protein	AJ245439	SA_map_164.for	AATCATCAAGTWCRTTACAYCATGG	52	164
			SA_map_1173.rev	CATTTACTCGAATTGTGTATGGTAC	52	1173
sdrC	Ser-Asp-rich fibrinogen-binding protein	AJ005645	SA_sdrC_149.for	GGCAGAACATACGAATGGAGAA	54	149
			SA_sdrC_689.for	ATTCTGTTAAAGAGGGCGATAC	54	689
			SA_sdrC_1604.rev	TTCATCTGTTGTCGTACGATCT	54	1604
sdrD	Ser-Asp-rich fibrinogen-binding protein	AJ005646	SA_sdrD_980.for	GACTGCAAAACATGATACTGCA	50	980
			SA_sdrD_2251.rev	TGTTACTGTAGTCGGTGTAATG	52	2251
			SA_sdrD_2897.rev	TCATTTCCAGAAGTTACTGAAG	50	2897
sdrE	Ser-Asp-rich fibrinogen-binding protein	AJ005647	SA_sdrE_803.for	AGTTGCACAACCAGCAGCAGTT	58	803
			SA_sdrE_2276.rev	GGCATCAGTAGTAGTTGTACCA	58	2276
spa	Immunoglobulin G-, A-, M-, and E-binding protein	M18264	SA_spa_729.for	ACTAGGTGTAGGTATTGCATCTGT	56	729
			SA_spa_1984.rev	TCCAGCTAATAACGCTGCACCTAA	56	1984

[a] Optimal annealing temperatures were determined by the success in performing PCR reactions with temperatures covering a range from 6°C below and above the primer-melting temperature.

for success and specificity, i.e., one clear band visible per PCR. All reactions with no amplification were repeated to exclude the possibility of technical errors. In addition, some of these reactions were repeated with newly extracted DNA and/or lowered annealing temperatures. However, none of the findings was altered by a second reaction. PCRs were then purified with the Montage PCR u96 kit (Millipore) according to the manufacturer's instructions.

Sequencing reactions were carried out with the Big Dye Terminator kit, version1.1 (Applied Biosystems), according to the manufacturer's instructions. Sequencing reaction cleanup was performed with the Montage Seq96 Sequencing Reaction Cleanup kit (Millipore), according to the manufacturer's instructions. Purified samples were analyzed with the ABI Prism 3100 Avant sequencer (Applied Biosystems), according to standard protocols.

Preparation of sequence data sets. First, alignments of all genes were created using ClustalW 1.81. Then, gaps introduced by ClustalW were, if needed, shifted to either side by one or two bases to produce sequences that are in frame. These alignments were taken for analysis of the genetic relationships between isolates. In a second procedure, indel sites were removed from alignments to produce data sets without gaps and that were in frame. Finally, for each gene a data set was compiled that comprised one copy of each allele. These were the data sets used to calculate nucleotide diversity and strength of purifying selection and to compare allelic and nucleotidic divergence.

Determination of sequence types. MLST was performed as previously described (11) using the *S. aureus* MLST database at http://www.mlst.net/ (Table 1). Sequences of unknown alleles were confirmed by repeating the procedure from a new DNA extraction of the isolate in question and were submitted to the MLST database. Furthermore, the obtained sequence types (STs) were assigned to previously identified clonal complexes (CCs) (CC1, CC5, CC8, CC22, CC30, and CC45) (12, 14).

***spa* typing.** The *spa* region C contains at its 3′ end the beginning of the first 24-bp *spa* repeat (39). Beginning with this repeat, Ridom *spa* allele numbers were determined at http://www.ridom.de/spaserver/. The concatenated allele numbers were taken to determine the Ridom *spa* type.

Molecular evolution. We calculated allelic diversity and nucleotide diversity (mean number of nucleotide differences between alleles) by using DNASP, version 4.00 (33). The proportion of synonymously (*ds*) and nonsynonymously (*dn*) mutated sites was determined using the implemented function in the software Start (37). *dn* and *ds* values were plotted against each other to produce data points in a graph. By measuring the strength of purifying selection (*dn*/*ds* ratio), the evolutionary forces of a gene are usually classified into three categories. Values of <1 indicate purifying selection, values approximately equal to 1 indicate balanced selection or neutral evolution, and values of >1 indicate diversifying selection or accelerated evolution.

Genetic relationship between isolates. For a reference and starting point, we calculated the gene tree of all MLST housekeeping genes. Gene sequences of all seven loci were concatenated to produce an in-frame sequence. To this, we compared gene trees of each gene. For representation of the other genomic compartments, we equally concatenated sequences of six core adhesion genes and four accessory adhesion genes. Maximum parsimony trees were calculated by using MEGA, version 3.0 (22). The statistical significance of branches was inferred by performing a bootstrap analysis with 1,000 repetitions.

Comparison of allelic and nucleotidic divergence. The underlying assumption of this analysis is that by comparing allelic and nucleotidic divergence it can be inferred whether clones in a population diversify predominantly by mutation which would result in a positive correlation between the two parameters or predominantly by recombination, which would result in a different relationship (14). Two data sets, based on either MLST housekeeping genes or core adhesion genes, were created. Pairwise comparisons of all STs were performed, and all pairs were sorted according to differences at one, two, three, four, five, six, and seven loci (up to six for core adhesion genes). For each pairing within a group, the number of nucleotide differences was then calculated per locus. The average number of differences was afterwards determined across loci. In this manner, a graph was produced where each group (*x* axis) was plotted against the average number of nucleotide differences across loci (*y* axis). The calculations were performed by using the BLUND program that was kindly provided by D. A. Robinson.

Epidemic and sporadic isolates. The rational of this comparison was to find out by means of molecular evolutionary analysis whether there was a difference related to either of the two groups of isolates. We aimed to look for each gene separately, whether or not there was a difference in nucleotide diversity and/or purifying selection between the two groups. A difference between the two groups could be imagined, due to the accumulation of mutations that occurred among sporadic isolates, which would be indicated by increased nucleotide diversity and/or reduced strength of purifying selection. We also aimed to examine whether epidemic isolates were grouped together in any of the obtained gene trees, thereby indicating a specific allele or pattern of alleles for epidemicity.

Nucleotide sequence accession numbers. Sequences were submitted to the EMBL database (accession numbers AM075836 to AM076337).

TABLE 3. Nucleotide sequence variation and purifying selection among alleles of 7 MLST housekeeping and 10 adhesion genes

Gene	Sequence length (bp)	No. of alleles	No. of polymorphic sites	Nucleotide diversity	Purifying selection (dn/ds)
Housekeeping genes					
arcC	456–457	6	8	0.0076	0.16
aroE	456–457	10	13	0.0112	0.26
glpF	465–466	4	8	0.0097	0.11
gmk	429	10	11	0.0103	0.11
pta	474	11	9	0.0071	0.20
tpi	402	8	9	0.0089	0.20
yqiL	516	10	13	0.0073	0.20
Core adhesion genes					
clfA	1,177	13	204	0.0681	0.21
clfB	1,196–1,198	9	136	0.0412	0.18
fnbA	1,381–1,390	19	538	0.1560	0.24
map	920–926	13	174	0.0739	0.17
sdrC	1,368–1,395	14	312	0.0936	0.17
spa	1,108–1,522	21	137	0.0260	0.12
Accessory adhesion genes					
ebpS	1,049–1,061	14	82	0.0285	0.59
fnbB	1,254–1,290	12	530	0.1779	0.31
sdrD	1,841	15	351	0.0729	0.24
sdrE	1,332	14	316	0.0740	0.25

RESULTS

Sequence analysis. Analysis of the 17 gene loci sequenced showed that, among the 58 MRSA isolates that cover the major clones found in Europe and the two MSSA isolates, a total of 24 strains did lack at least 1 gene locus. One strain lacked *ebpS*, 11 strains lacked *fnbB*, 1 strain lacked *sdrD* only, 6 strains lacked *sdrE* only, and 6 strains lacked both *sdrD* and *sdrE*. Therefore, these four genes were considered "accessory" for further analysis. In contrast, *clfA*, *clfB*, *fnbA*, *map*, *sdrC*, and *spa* were present in all strains and were classified as "core" adhesion genes (Table 1).

Indel mutations were observed among alleles of genes *clfB*, *ebpS*, *fnbA*, *fnbB*, *map*, *sdrC*, and *spa*. The great majority of these did not lead to a frameshift and produced a continuous open reading frame. Four alleles (*arcC*, isolate 55; *clfB*, isolates 25 and 40; and *spa*, isolate 31) comprised one or two base insertions. This led to appearance of early stop codons in the open reading frame. Thus, frameshift mutations occurred among one MLST housekeeping gene and two core adhesion genes. In addition, we found three new alleles (*aroE*, isolate 45; *glpF*, isolate 62; and *gmk*, isolate 33) that have not yet been submitted to the MLST database.

Our collection comprised isolates of the same clone that were sampled either at different locations or at different moments in time. For five of the clones, isolates showed the same genotype with all markers used in this study. For the remaining four clones, differences were observed (Table 1). For example, two isolates of EMRSA-15 from Portugal differed at locus *gmk* (one comprised allele 5, the other allele 33). This indicated either that two different strains of EMRSA-15 were introduced to Portugal or that one strain accumulated a mutation during its dissemination.

Determination of sequence types. Based on the seven MLST housekeeping genes, the complete data set with all 60 MRSA isolates contained 24 STs. BURST analysis of all 60 isolates showed that there were six distinct groups. Nineteen STs were within the five major MRSA lineages that were previously described (CC5, CC8, CC22, CC30, and CC45). One additional group was formed by the identical ST of MSSA476, community-acquired MRSA MW2, and one new ST. This group corresponded to the previously described CC1. The remaining two STs, isolates 44 and 90, were singletons and were new STs that are not yet described in the MLST database.

Molecular evolution. Variable numbers of alleles were observed for all genes (Table 3). In general, there were more adhesion gene alleles than housekeeping gene alleles, which can be expected due to the greater length of adhesion gene alleles. Calculation of nucleotide diversity between alleles takes into account the different sizes of the products and thus allows the comparison of values between genes and classes of genes. According to the classification of genes, we found that values for core adhesion genes (mean = 0.077 ± 0.046) or accessory adhesion genes (mean = 0.088 ± 0.063) were significantly greater than those for MLST housekeeping genes (mean = 0.009 ± 0.002) ($P = 0.0023$ and $P = 0.007$, respectively), meaning that there was more nucleotide diversity between adhesion genes than between housekeeping genes (Table 3). The amount of nucleotide diversity ranged from double for *spa* to >10 times as much for *fnbA* and *fnbB*.

The strength of purifying selection was then calculated as *dn*/*ds*, and showed that variation was high between genes and classes of genes. When tested for significance, there was no

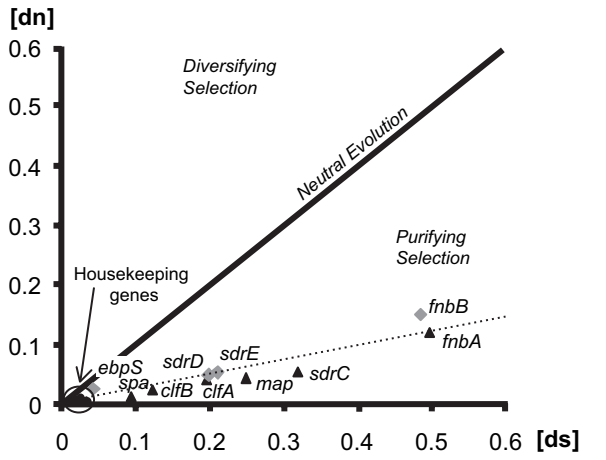

FIG. 1. Purifying selection. The aligned sequence data of each gene were taken to calculate the proportion of synonymous (*ds*) and nonsynonymous (*dn*) substitutions. The plotted values of each gene are shown. The ratio of *dn*/*ds* indicates purifying selection (positive Darwinian selection) if values are <1, diversifying selection if values are >1, and balancing selection or neutral evolution if values are close to 1. Three classes of genes are represented: housekeeping genes *arcC*, *aroE*, *glpF*, *gmk*, *pta*, *tpi*, and *yqiL* (circles); core adhesion genes (triangles); and accessory adhesion genes (diamonds). Gene names of the latter two classes are shown in the graph. Data points of all genes are located in the area of purifying selection. The dotted line shows the regression line calculated on the basis of all data points ($R^2 = 0.9151$).

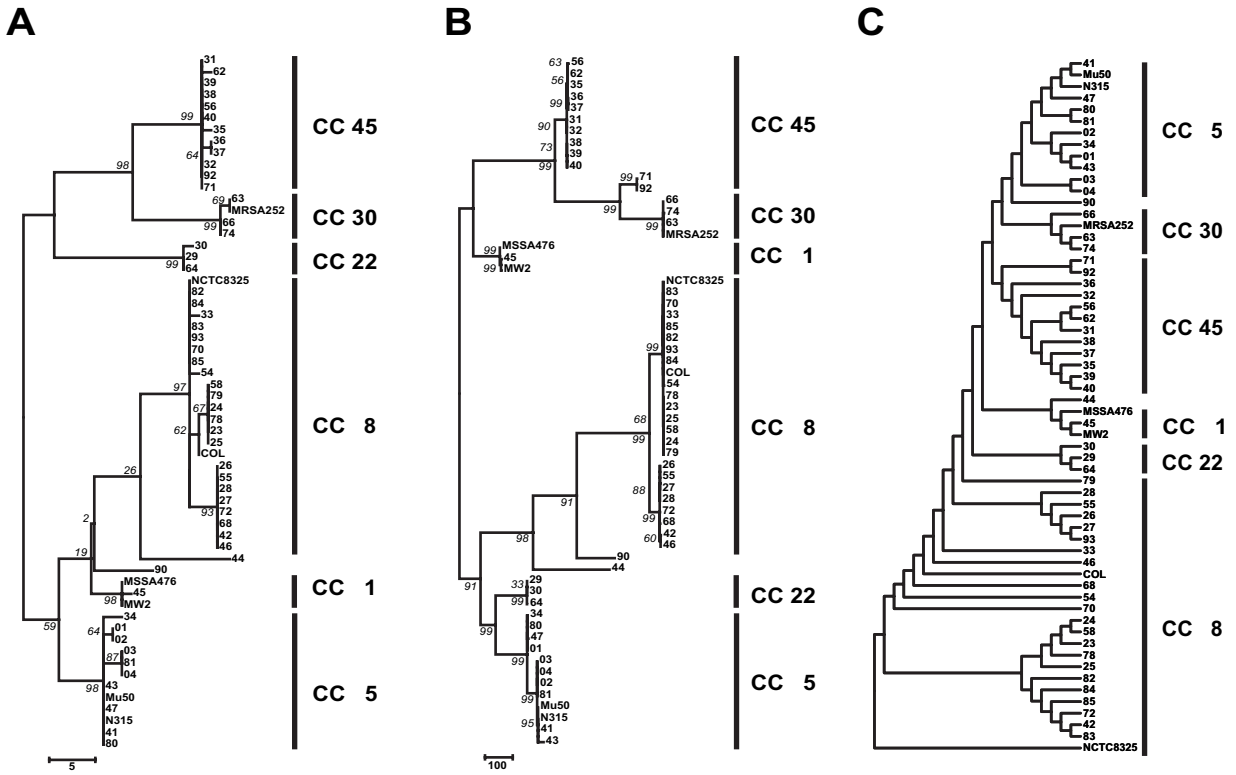

FIG. 2. Genetic relationship between isolates. The sequences obtained from each isolate for each gene were taken to produce alignments without gaps. We then concatenated separately the gene sequences to produce data sets for seven MLST housekeeping genes (A), six core adhesion genes (B), and four accessory adhesion genes (C). Maximum parsimony trees were constructed for all three data sets with MEGA 3.0. For the data shown in panels A and B, the branch lengths were also calculated and their scales are indicated at the lower left of each panel. Bootstrap values were calculated for 1,000 repetitions and are indicated at branches. In addition, STs were determined for all isolates based on housekeeping genes using the MLST database, and these were used to assign each ST to published CCs. CCs are indicated in all three panels. We observed that all groups of CCs were resolved in each concatenated gene tree.

significant difference in strength of purifying selection between the MLST housekeeping genes (mean $dn/ds = 0.176 \pm 0.054$) and core adhesion genes (mean $dn/ds = 0.18 \pm 0.041$) ($P = 0.879$). However, when each of these two classes was compared to the class of accessory adhesion genes (mean $dn/ds = 0.347 \pm 0.165$), there was less purifying selection in the latter class of genes, with weak significance ($P = 0.0287$ and $P = 0.0408$, respectively).

Values of *dn* and *ds* of all genes were plotted in a graph (Fig. 1). All genes were subject to purifying selection and were therefore found in the bottom part of the plot. A regression line with a good correlation ($R^2 = 0.9151$) could be calculated on the basis of all values. MLST housekeeping genes had very low values of *dn* and *ds*, due to their low nucleotide diversity. They grouped together at the lower-left-hand part of the plot. Synonymous and nonsynonymous substitution rates of adhesion genes were more variable between genes, as was previously indicated by the variation of nucleotide diversity. The weaker purifying selection in accessory adhesion genes was indicated in the graph by the fact that these genes were all plotted above the regression line (Fig. 1).

Genetic relationship between isolates. The concatenated set of MLST housekeeping genes (Fig. 2A) produced a gene tree with weakly supported groups of isolates (low bootstrap values). However, all five groups of isolates known as major MRSA CCs were identified.

The concatenated set of core adhesion genes produced a gene tree with very well-supported groups of isolates (Fig. 2B). High bootstrap values were found at nodes that separated not only the major groups of clonal complexes but also groups previously separated by multiprimer randomly amplified polymorphic DNA (RAPD) analysis. In CC8, all isolates of ST239 were well grouped together. This group was previously identified as 2a by multiprimer RAPD (4). Within CC45, two isolates (71 and 92) were distinctly separated and corresponded to RAPD group 6 (4).

The reconstruction of the accessory adhesion gene tree differed from the other gene trees in that calculation of branch lengths was impossible because there were no sites in common in the concatenated data set. Therefore, only the stepwise differences between isolates based on maximum parsimony were shown (Fig. 2C). Still, all major MRSA complexes were found to be concordant in the grouping of the tree.

Visibly, there was high congruence between the gene trees of the different classes of concatenated data sets, indicating that recombination had little effect on the population structure. When individual gene trees were constructed and compared, we also observed a high congruence between trees (data not

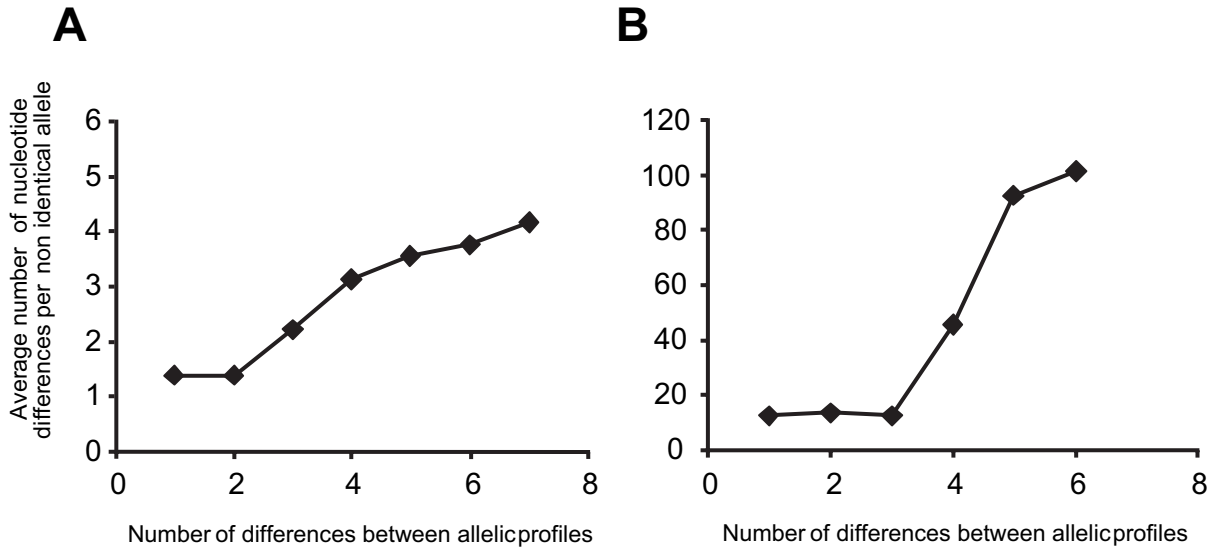

FIG. 3. Sequence diversity versus allelic diversity. Two data sets were analyzed separately: MLST housekeeping genes (A) and core adhesion genes (B). They comprised 24 and 34 STs, respectively. Pairwise comparisons of all STs were performed. They were grouped according to differences at one, two, three, four, five, six, and seven loci (up to six for core adhesion genes). Within each group, the number of nucleotide differences was then calculated for each locus separately; the average was afterwards determined across loci. In this way, a graph was produced where each group (*x* axis) was plotted against the average number of nucleotide differences across loci (*y* axis). The calculations were performed using the BLUND program that was kindly provided by D. A. Robinson. We observed a positive trend for MLST housekeeping genes, indicating that diversification was mainly due to accumulation of mutations. There was also a positive trend from three to six different loci among core adhesion genes. However, there was no positive correlation for up to three different loci among core adhesion genes; this could be explained by recombination.

shown), while the best congruence was observed with *clfA* and the least was observed with *clfB*.

Evidence for homologous exchange. Gene transfer was suspected when the same adhesion gene allele was found in isolates from distant phylogenetic lineages (CCs) (Table 1). This way, four events of homologous exchange were identified. For the gene *clfB*, allele 1 was found in some isolates of CC8 and in all isolates of CC5; allele 6 was found in the remaining isolates of CC8 and in all isolates of CC30. For the gene *fnbA*, allele 13 was found in isolates of both CC1 and CC22. For the gene *spa*, allele 8 was found in some isolates of CC8 and one isolate of CC30. We emphasize that the *spa* gene tree must be taken with care; due to many indel mutations throughout the whole sequence, only very few repeats and domains were taken to construct an alignment. Thus, there was potential for error. Nucleotide diversification could have occurred after an event of gene transfer. In this case, if highly similar alleles were found in different phylogenetic lineages, four other events of transfer could be suspected (data not shown).

Comparison of allelic and nucleotidic divergence. The goal of this analysis was to identify whether MLST housekeeping and core adhesion genes diversified the same way. All 24 STs of MLST housekeeping genes and all 34 STs of the core adhesion genes were taken for analysis. The nucleotidic divergence was >10 times greater among adhesion genes than among MLST housekeeping genes, which was expected due to the differences in nucleotide diversity (Fig. 3). To be able to observe differences in the shape of the curve, we plotted each data set independently. Both plots clearly showed a positive correlation, indicating accumulation of de novo mutations. However, a small difference was observed between the two curves, in that the slope was null for up to two allelic differences for MLST housekeeping genes, whereas it was null for up to three allelic differences for core adhesion genes.

Epidemic and sporadic isolates. Nucleotide diversity and purifying selection were similar between epidemic and sporadic isolates (data not shown). Also, neither any of the single-gene trees nor a concatenated gene tree showed a specific grouping of epidemic or sporadic isolates (Fig. 2). Thus, there was no specific allele (genetic element) or pattern of alleles (genetic signature) of epidemic or sporadic isolates.

DISCUSSION

The sequence data of 17 genes from a diverse collection of 58 MRSA and 2 MSSA isolates were used to study the evolution of the three well-defined groups of genes that represent conceptually different parts of the *S. aureus* genome. Based on their continuous or discontinuous presence in our collection of strains, we classified the 10 adhesion genes into two groups, core adhesion genes and accessory adhesion genes, next to the housekeeping MLST genes. These genes represented a cross section of the *S. aureus* genome from the slowly evolving genomic core (MLST housekeeping genes) to the highly polymorphic accessory part of the genome (accessory adhesion genes).

Molecular evolution. High degrees of polymorphism were found among the alleles of adhesion genes. The great number of nucleotide differences was expected, because cell surface genes are generally considered to be highly polymorphic and because a similar finding was previously reported for *sas* genes, a group of genes that carried the LPXTG protein motif for

exportation to the cell surface (32). The studied *sas* genes were located in one region of the *S. aureus* chromosome, which (at least for strains of CC8) was subject to a particular evolution comprising homologous exchange (31). The 10 adhesion genes were distributed on seven different places over the entire chromosome; thus, we are confident that the great number of nucleotide differences is a general characteristic of cell surface genes in *S. aureus*.

There was reason to suspect that one of the mechanisms to create nucleotide differences among adhesion genes would be diversifying selection. However, in our data set, genes of all classes showed similar levels of purifying selection. Only among accessory adhesion genes, purifying selection was a little less strong than in the other classes. This indicates that functionality of adhesion genes is very important and that despite the great divergence between alleles there is no or little accumulation of deleterious mutations. In fact, the *dn*/*ds* plot of all genes followed a remarkable linear regression. If no indication of diversifying selection is observed in cell surface genes, which are likely the most polymorphic, one could suspect this is true for all genes in the *S. aureus* genome.

Genetic relationship. The three concatenated gene trees reconstructed from core or accessory genome genes revealed a remarkably high congruence to previously published data. They were highly congruent to CCs that were found in another study using BURST on MLST housekeeping genes (12), and they were also in agreement with a multiprimer RAPD approach where 14 of the 60 *S. aureus* isolates were already analyzed (4). This means that adhesion genes followed the clonal evolution that was previously described for MLST housekeeping genes.

One of the problems using MLST housekeeping genes in phylogenetic approaches is the small number of nucleotide differences between alleles. Therefore, only analyses that were based on the concatenated sequences of all seven genes led to resolution of groups that are congruent with CCs. Increased divergence between adhesion gene alleles offered the possibility to construct groupings with higher bootstrap values. These results support the hypothesis that CCs are significant genetic entities within the *S. aureus* population.

In this study, it was possible to correctly classify all isolates in previously defined CCs with only one marker, *clfA*. It was even possible to separate the two groups of strains within CC8 which were the result of gene transfer (31), thereby showing that their separation-specific mutations have already occurred in each lineage. This renders it a good candidate for future typing approaches.

We would like to emphasize that *clfB*, which is located at a different genomic position than *clfA*, was the least congruent of all studied genes. This was in contrast to an earlier study of adhesion genes, where *clfB* was proposed as potentially useful for phylogenetic analyses because it showed highest congruence to the expected groups of isolates (18). We think that this difference could be because the previous study was based on another part of *clfB* that contained various numbers of repeats. Such a region is difficult to analyze phylogenetically, and its evolution might differ from the rest of the gene.

Evidence for homologous exchange. As had been shown in previous studies, we observed the same alleles for *clfB* and *spa* in isolates of CC30 and of ST239 (CC8), indicating gene transfer between both lineages. This transfer has already been described previously with the passage of >200 kb of DNA (which includes the genes *arcC*, *clfB*, and *spa*). In addition, we observed an as-yet-unreported transfer where isolates of CC8 were characterized by a *clfB* allele which was shared by all isolates of CC5, suggesting also the passage of this allele from CC5 to CC8. Taken together, all alleles of CC8 for *clfB* were acquired from other CCs, while the ancestral allele could not be found in our collection of isolates. Further studies will need to be performed to know whether this is true for the whole population and whether the ancestral allele was irreversibly lost.

To our surprise, the gene *arcC*, which on all sequenced *S. aureus* genomes is located directly next to *clfB*, was not shown to be part of this transfer. It was previously argued that the nonclonal evolutionary history of *arcC* might be a "hitchhiking effect," due to the physical proximity to *clfB*, which was predicted to undergo a different evolution (14). This homologous exchange is at odds with the hitchhiking suggestion, but it might be an exception. Further studies are needed to know whether or not hitchhiking has greatly contributed to the evolution of *arcC*.

The overall relatively small amount of evidence for homologous exchange is in contrast to a previous study that analyzed a set of accessory virulence genes and suspected frequent events of transfer (28). There, evidence for gene transfer was considered, due to the observed presence and absence of accessory virulence genes within the same phylogenetic lineage. We think that such a pattern could still be explained with a predominantly clonal evolution, where genes were subsequently lost from certain isolates within lineages; such a pattern does not in itself account as evidence for homologous exchange. Such a model was already proposed for the molecular evolution of an accessory pathogenicity island (RD13), where the consecutive loss of genes was the most parsimonious model of evolution, indicating a predominantly clonal evolution (15).

Because we included the most variable class of genes, because these were distributed over the whole chromosome, and because we analyzed a highly diverse collection of isolates, we have maximized the chances to find evidence for homologous exchange. Therefore, we think that the high congruence between gene trees as presented in this study completed a panel of evidence for a predominantly clonal evolution of *S. aureus* (14, 15, 32).

Comparison of allelic and nucleotidic divergence. The graph produced by MLST housekeeping gene analysis showed a clear linear and positive trend between nucleotidic and allelic divergence, suggesting a predominantly clonal evolution. It confirmed the analysis of a previous study using a more extensive MLST gene data set obtained from a different collection of isolates (14) and thus confirmed the robustness of our data. The same general trend was observed with adhesion genes. The fact that a predominantly clonal evolution was observed in both conserved housekeeping genes and highly polymorphic cell surface genes suggests that the majority of *S. aureus* genes followed this evolution.

Surprisingly, a more detailed look at both graphs revealed that for small numbers of allelic differences (up to two in the case of MLST housekeeping genes and up to three in the case

of core adhesion genes), there was no clear positive correlation with nucleotidic diversity. This was also true for the previous analysis but did not receive attention (14). No correlation is generally perceived as the result of recombination that overcomes the effects of divergence due to de novo point mutations. Our data suggest that recombination occurs between isolates differing by only few alleles. Thus, despite a predominantly clonal population structure of *S. aureus*, we postulate that recombination occurs more frequently between closely related strains than between distant phylogenetic lineages. Such model of a population structure was proposed by M. Tibayrenc and was called "strict homogamy" (38). This model supposed the existence of "discrete genetic entities" within a species that can be seen as isolated populations between which no gene flow occurs. Importantly, we should note that CCs generally contain pairs of strains that differ by no more than three loci, which would indicate that a clonal complex delimits an isolated population.

Biological significance. The attachment of a bacterial cell to the host is one fundamental step in the line of events for host infection and thus could contribute to epidemicity of MRSA strains. One of the reasons that we found no evidence for genetic differences between epidemic and sporadic isolates might be the limited number of adhesion genes used in this study. However, the analysis of seven other *S. aureus* surface genes (*sas*) also showed high congruence with MLST housekeeping genes, supporting the hypothesis that there are no phylogenetic differences between epidemic and sporadic isolates. Thus, it seems that there is no link or correlation between epidemic potential and a certain adhesion gene or allele. However, it could also be imagined that only localized parts of a protein are linked to epidemic behavior, resulting in a specific nucleotide signature within genes. This has not yet been addressed. Also, due to the likely homogamic population structure, it is possible that epidemic clones have independently evolved in each CC and cannot be compared between CCs. Indeed, an approach including only strains of CC8 isolated in the same country clearly showed that epidemic strains (Brazilian clone) displayed significant physiological differences to other strains of the same CC. The epidemic strains produced more biofilm and showed stronger adherence to epithelial cells (1).

Concluding remarks. *S. aureus* adhesion genes can be regarded as highly polymorphic and, compared to housekeeping genes, represent a highly variable genomic compartment. We showed that purifying selection was present among all genes, whereas it was less strong among accessory adhesion genes. High congruence between gene trees of different genomic compartments indicated a clonal population structure and showed that it is remarkably prevalent in *S. aureus* genome evolution. Therefore, the association of a specific genetic element with epidemic MRSA clones failed. However, evidence for homologous exchange was detected, and we showed a previously undescribed transfer between CC5 and CC8. We also found that closely related *S. aureus* isolates that fit the demarcation of CCs showed evidence of recombination. This suggests that despite the remarkably overall clonal *S. aureus* population structure, recombination might significantly contribute to evolution of CCs.

ACKNOWLEDGMENTS

We thank H. de Lencastre, M. Aires de Sousa, K. W. Larssen, W. Wannet, M. Struelens, J. Etienne, and F. Vandenesch for providing us with epidemic and sporadic MRSA strains. We are very grateful to D. A. Robinson for kindly providing the BLUND program and for helpful online discussions.

This work was funded with grant 3200B0-101044 from the Swiss National Foundation for Research.

REFERENCES

1. **Amaral, M. M., L. R. Coelho, R. P. Flores, R. R. Souza, M. C. Silva-Carvalho, L. A. Teixeira, B. T. Ferrerira-Carvalho, and A. M. S. Figueiredo.** 2005. The predominant variant of the Brazilian epidemic clonal complex of methicillin-resistant *Staphylococcus aureus* has an enhanced ability to produce biofilm and to adhere to and invade airway epithelial cells. J. Infect. Dis. **192:**801–810.
2. **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.
3. **Betley, M. J., and J. J. Mekalanos.** 1985. Staphylococcal enterotoxin A is encoded by phage. Science **229:**185–187.
4. **Blanc, D. S., A. L. Banuls, P. M. Hauser, P. Moreillon, P. Francioli, M. Tibayrenc, and the Swiss MRSA Group.** 2000. Methicillin-resistant *Staphylococcus aureus*: phylogenetic relatedness between European epidemic clones and Swiss sporadic strains. Microb. Drug Resist. **6:**231–238.
5. **Blanc, D. S., P. Francioli, A. Le Coustumier, L. Gazagne, E. Lecaillon, P. Gueudet, F. Vandenesch, and J. Etienne.** 2001. Reemergence of gentamicin-susceptible strains of methicillin-resistant *Staphylococcus aureus* in France: a phylogenetic approach. J. Clin. Microbiol. **39:**2287–2290.
6. **Blanc, D. S., C. Petignat, P. Moreillon, J. M. Entenza, M. C. Eisenring, H. Kleiber, A. Wenger, N. Troillet, C. H. Blanc, and P. Francioli.** 1999. Unusual spread of a penicillin-susceptible methicillin-resistant *Staphylococcus aureus* clone in a geographic area of low incidence. Clin. Infect. Dis. **29:**1512–1518.
7. **Deplano, A., W. Witte, W. J. Van Leeuwen, Y. Brun, and M. J. Struelens.** 2000. Clonal dissemination of epidemic methicillin-resistant *Staphylococcus aureus* in Belgium and neighboring countries. Clin. Microbiol. Infect. **6:**239–245.
8. **de Sousa, M. A., I. S. Sanches, M. L. Ferro, M. J. Vaz, Z. Saraiva, T. Tendeiro, J. Serra, and H. de Lencastre.** 1998. Intercontinental spread of a multidrug-resistant methicillin-resistant *Staphylococcus aureus* clone. J. Clin. Microbiol. **36:**2590–2596.
9. **de Sousa, M. A., and H. de Lencastre.** 2003. Evolution of sporadic isolates of methicillin-resistant *Staphylococcus aureus* (MRSA) in hospitals and their similarities to isolates of community-acquired MRSA. J. Clin. Microbiol. **41:**3806–3815.
10. **Elphinstone, M. S., G. N. Hinten, M. J. Anderson, and C. J. Nock.** 2003. An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. Mol. Ecol. Notes **3:**317–320.
11. **Enright, M. C., N. P. J. 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.
12. **Enright, M. C., D. A. Robinson, G. Randle, E. J. Feil, H. Grundmann, and B. G. Spratt.** 2002. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). Proc. Natl. Acad. Sci. USA **99:**7687–7692.
13. **Feil, E. J.** 2004. Small change: keeping pace with microevolution. Nat. Rev. Microbiol. **2:**483–495.
14. **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. J. Day.** 2003. How clonal is *Staphylococcus aureus*? J. Bacteriol. **185:**3307–3316.
15. **Fitzgerald, J. R., S. D. Reid, E. Ruotsalainen, T. J. Tripp, M. Y. Liu, R. Cole, P. Kuusela, P. M. Schlievert, A. Jarvinen, and J. M. Musser.** 2003. Genome diversification in *Staphylococcus aureus*: molecular evolution of a highly variable chromosomal region encoding the staphylococcal exotoxin-like family of proteins. Infect. Immun. **71:**2827–2838.
16. **Fitzgerald, J. R., D. E. Sturdevant, S. M. Mackie, S. R. Gill, and J. M. Musser.** 2001. Evolutionary genomics of *Staphylococcus aureus*: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. Proc. Natl. Acad. Sci. USA **98:**8821–8826.
17. **Gill, S. R., D. E. Fouts, G. L. Archer, E. F. Mongodin, R. T. DeBoy, J. Ravel, I. T. Paulsen, J. F. Kolonay, L. Brinkac, M. Beanan, R. J. Dodson, S. C. Daugherty, R. Madupu, S. V. Angiuoli, A. S. Durkin, D. H. Haft, J. Vamathevan, H. Khouri, T. Utterback, C. Lee, G. Dimitrov, L. X. Jiang, H. Y. Qin, J. Weidman, K. Tran, K. Kang, I. R. Hance, K. E. Nelson, and C. M. Fraser.** 2005. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. J. Bacteriol. **187:**2426–2438.

18. **Gomes, A. R., S. Vinga, M. Zavolan, and H. de Lencastre.** 2005. Analysis of the genetic variability of virulence-related loci in epidemic clones of methicillin-resistant *Staphylococcus aureus*. Antimicrob. Agents Chemother. **49:** 366–379.
19. **Holden, M. T. G., E. J. Feil, J. A. Lindsay, S. J. Peacock, N. P. J. 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. Feltwell, 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. Rabbinowitsch, 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.
20. **Iandolo, J. J., V. Worrell, K. H. Groicher, Y. D. Qian, R. Y. Tian, S. Kenton, A. Dorman, H. G. Ji, S. P. Lin, P. Loh, S. L. Qi, H. Zhu, and B. A. Roe.** 2002. Comparative analysis of the genomes of the temperate bacteriophages phi 11, phi 12 and phi 13 of *Staphylococcus aureus* 8325. Gene **289:**109–118.
21. **Koreen, L., S. V. Ramaswamy, E. A. Graviss, S. Naidich, J. A. Musser, and B. N. Kreiswirth.** 2004. *spa* typing method for discriminating among *Staphylococcus aureus* isolates: Implications for use of a single marker to detect genetic micro- and macrovariation. J. Clin. Microbiol. **42:**792–799.
22. **Kumar, S., K. Tamura, and M. Nei.** 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinformatics **5:**150–163.
23. **Kuroda, M., T. Ohta, I. Uchiyama, T. Baba, H. Yuzawa, I. Kobayashi, L. Z. Cui, A. Oguchi, K. Aoki, Y. Nagai, J. Q. 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 meticillin-resistant *Staphylococcus aureus*. Lancet **357:**1225–1240.
24. **Lan, R., and P. Reeves.** 2000. Intraspecies variation in bacterial genomes: the need for a species genome concept. Trends Microbiol. **8:**396–401.
25. **Larssen, K. W., G. Iversen, T. Jacobsen, K. Bergh, and O. Scheel.** 2004. MRSA outbreaks in nursing homes in central Norway. Clin. Microbiol. Infect. **10:**126.
26. **Lindsay, J. A., A. Ruzin, H. F. Ross, N. Kurepina, and R. P. Novick.** 1998. The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus*. Mol. Microbiol. **29:**527–543.
27. **Maiden, M. C. J., J. A. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman, and B. G. Spratt.** 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. USA **95:**3140–3145.
28. **Moore, P. C. L., and J. A. Lindsay.** 2001. Genetic variation among hospital isolates of methicillin-sensitive *Staphylococcus aureus*: evidence for horizontal transfer of virulence genes. J. Clin. Microbiol. **39:**2760–2767.
29. **Oliveira, D., R. Santos-Sanches, M. Mato, G. Tamayo, D. Ribeiro, D. Costa, and H. de Lencastre.** 1998. Virtually all methicillin-resistant *Staphylococcus aureus* (MRSA) infections in the largest Portuguese teaching hospital are caused by two internationally spread multiresistant strains: the 'Iberian' and the 'Brazilian' clones of MRSA. Clin. Microbiol. Infect. **4:**373–384.
30. **Peacock, J. E., D. R. Moorman, R. P. Wenzel, and G. L. Mandell.** 1981. Methicillin-resistant *Staphylococcus aureus*—microbiologic characteristics, antimicrobial susceptibilities, and assessment of virulence of an epidemic strain. J. Infect. Dis. **144:**575–582.
31. **Robinson, D. A., and M. C. Enright.** 2004. Evolution of *Staphylococcus aureus* by large chromosomal replacements. J. Bacteriol. **186:**1060–1064.
32. **Robinson, D. A., and M. C. Enright.** 2003. Evolutionary models of the emergence of methicillin-resistant *Staphylococcus aureus*. Antimicrob. Agents Chemother. **47:**3926–3934.
33. **Rozas, J., and R. Rozas.** 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics **15:**174–175.
34. **Sanches, I. S., M. Aires de Sousa, L. Sobral, I. Calheiros, L. Felicio, I. Pedra, and H. de Lencastre.** 1995. Multidrug-resistant Iberian epidemic clone of methicillin-resistant *Staphylococcus aureus* endemic in a hospital in northern Portugal. Microb. Drug Resist. **1:**299–306.
35. **Sanches, I. S., M. Ramirez, H. Troni, M. Abecassis, M. Padua, A. Tomasz, and H. Delencastre.** 1995. Evidence for the geographic spread of a methicillin-resistant *Staphylococcus aureus* clone between Portugal and Spain. J. Clin. Microbiol. **33:**1243–1246.
36. **Shopsin, B., M. Gomez, S. O. Montgomery, D. H. Smith, M. Waddington, D. E. Dodge, D. A. Bost, M. Riehman, S. Naidich, and B. N. Kreiswirth.** 1999. Evaluation of protein A gene polymorphic region DNA sequencing for typing of *Staphylococcus aureus* strains. J. Clin. Microbiol. **37:**3556–3563.
37. **Smith, J. M.** 1999. The detection and measurement of recombination from sequence data. Genetics **153:**1021–1027.
38. **Tibayrenc, M.** 1995. Population genetics of parasitic protozoa and other microorganisms. Adv. Parasitol. **36:**47–115.
39. **Uhlén, M., B. Guss, B. Nilsson, S. Gatenbeck, L. Philipson, and M. Lindberg.** 1984. Complete sequence of the staphylococcal gene encoding protein A: a gene evolved through multiple duplications. J. Biol. Chem. **259:**1695–1702.
40. **Wannet, W. J. B., E. Spalburg, M. Heck, G. N. Pluister, R. J. L. Willems, and A. J. de Neeling.** 2004. Widespread dissemination in The Netherlands of the epidemic Berlin methicillin-resistant *Staphylococcus aureus* clone with low-level resistance to oxacillin. J. Clin. Microbiol. **42:**3077–3082.
41. **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.