Structural Variation of Wall Teichoic Acid affects Colonization, Virulence, and Evolution of *Staphylococcus epidermidis*

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

Vorgelegt von

Dr. med. Xin Du

Aus Zhejiang, China

Tübingen 2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	08.11.2019
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Andreas Peschel
2. Berichterstatter:	Prof. Dr. Friedrich Götz

Table of contents

To my family

Abstract

The structures of wall teichoic acids (WTAs), the most abundant component on the cell surface of gram-positive bacteria, are highly diverse. WTAs act as the specific receptors for phages. *Staphylococcus epidermidis* is a commensal bacterium found on almost all human beings. Some clonal lineages of *S. epidermidis* are frequently causing important healthcare-associated infections. However, the mechanism of how these lineages are pathogenic remains unknown. Besides, *S. epidermidis* has a high resistance rate to antibiotics since most strains carry SCC*mec* which might be transferred by phages from *Staphylococcus aureus*. This raises a long-existing question of how species-specific phages transfer mobile genetic elements from *S. epidermidis* to *S. aureus* or vice versa. This Ph.D. project found that some *S. epidermidis* lineages isolated from sites of infection presented an additional WTA structure similar to *S. aureus*, allowing infection by *S. aureus* phages thus promoting gene sharing between *S. epidermidis* and *S. aureus* via horizontal gene transfer (HGT). A novel gene cluster, *tarIJLM2*, is responsible for the expression of this *S. aureus*-type WTA on *S. epidermidis*. Moreover, this WTA structural variation contributes to the virulence of bacteria, which indicates that it is an effective biomarker for detecting pathogenic strains of *S. epidermidis*.

Thus, WTAs could control the direction of phage-mediated HGT among *staphylococci* by affecting interactions with different phages. Special WTA structures could also determine a certain bacterial lineage to be commensal or invasive. The study of WTA structural variations and the discovery of novel phages not only pave the way to elucidate the mechanism of s*taphylococci* evolution but also provide new targets for antibiotics and vaccines fighting against pathogenic s*taphylococci*.

Zusammenfassung

Die Strukturen der Wand-Teichonsäuren (WTAs), der am häufigsten vorkommenden Komponente auf der Zelloberfläche von Gram-positiven Bakterien, weisen eine hohe Diversität auf. Die WTAs fungieren als spezielle Rezeptoren für die Phagen. *Staphylococcus epidermidis* ist als kommensales Bakterium sowohl auf der Haut als auch in der Nase fast aller Menschen zu finden. Bestimmte klonale Linien von *S. epidermidis* verursachen häufig schwere nokosomiale Infektionen. Die Ursachen dafür, dass sich diese Linien besonders pathogen verhalten, sind jedoch immer noch weitgehend ungeklärt. Ein weiteres wichtiges Merkmal von *S. epidermidis* ist die hohe Resistenzrate gegen Antibiotika, da die meisten *S. epidermidis*-Stämme das SCC*mec*-Element tragen, das mittels Phagen von *Staphylococcus aureus* übertragen wird. Daher stellt sich die Frage, wie die artspezifischen Phagen den Transfer mobiler genetischer Elemente von *S. aureus* nach *S. epidermidis* oder umgekehrt bewirken. Diese Arbeit zeigt, dass einige *S. epidermidis*-Linien, isoliert aus Infektionen, eine zusätzliche WTA-Polymerstruktur aufweisen, die der von *S. aureus* ähnelt. Diese WTA vom *S. aureus*-Typ ermöglicht eine Infektion durch *S. aureus*-Phagen und fördert somit die gemeinsame Nutzung von Genen zwischen *S. epidermidis* und *S. aureus* mittels Gene horizontal (HGT). Ein bislang unbeschriebenes Gencluster, *tarIJLM2,* ist für die Expression von WTA des *S. aureus*-Typs in *S. epidermidis* verantwortlich. Darüber hinaus trägt diese WTA-Strukturvariation zur Virulenz der Bakterien bei, was darauf hinweist, dass *tarIJLM2* ein hilfreicher Biomarker für den Nachweis pathogener *S. epidermidis*-Stämme sein könnte.

Daher könnte WTA die die Richtung des Phagen-vermittelten HGT zwischen Staphylokokken kontrollieren, indem die Interaktion mit verschiedenen Phagen beeinflusst wird. Besondere WTA-Strukturen könnten entscheiden, ob bestimmte Bakterien sich kommensal oder invasiv verhalten. Die Untersuchung von WTA-Strukturvarianten und die Entdeckung von neuen Phagen ebnet nicht nur den Weg zum Verständnis der Evolution von Staphylokokken, sondern führt auch zur Identifizierung von neuen Angriffspunkten von Antibioka und Impfstoffen, die gegen pathogene Staphylokokken gereichtet sind.

Chapter 1

-

General Introduction

Staphylococcus epidermidis

Staphylococcus epidermidis belongs to Staphylococcus, which is a genus of gram-positive bacteria in the family *Staphylococcaceae*, where 'staphylo-' means 'bunch of grapes' and '-coccus' means 'grain, seed, berry', observations based on its appearance under a microscope. *Staphylococci* are common commensal bacteria found in humans. Among all Staphylococcus species, *S. epidermidis* is the most frequently isolated species from human skin and mucous membranes. As with other Staphylococcus, *S. epidermidis* can grow in both aerobic and anaerobic environments, contains a cell wall structure including peptidoglycan and teichoic acid, and has a genomic DNA GC content in the range of 30-40%.

S. epidermidis is distinguished from the more virulent *Staphylococcus aureus* because it lacks the enzyme coagulase that causes aggregation of human serum. Thus, *S. epidermidis* clustered into the group of coagulase-negative *staphylococci* (CoNS). *S. epidermidis*, together with *Staphylococcus caprae*, *Staphylococcus capitis*, and *Staphylococcus saccharolyticus* belongs to the epidermidis cluster group, which is most closely related to *Staphylococcus haemolyticus*[1]. *S. epidermidis* has recently gained increasing attention from researchers because it is one of the most important pathogens causing nosocomial diseases in the hospital. However, *S. epidermidis* mainly causes diseases that are not life threatening. The widespread and difficult-to-treat characteristics make *S. epidermidis* a serious burden for our society. Compared to the well-studied *S. aureus*, more research on *S. epidermidis* is urgently needed to better develop novel prevention and therapeutic strategies.

***S. epidermidis* as commensal bacteria**

The nose is the first organ of the human respiratory system. The nasal cavity makes many contacts with the external environment, which is inhabited by a wide variety of commensal and

pathogenic bacteria. While *S. aureus* as a commensal bacterium colonizes approximately 30% of human anterior nares, *S. epidermidis* colonizes in this niche of almost all human individuals (>90%)[2,3]. *S. epidermidis* can compete against and displace *S. aureus* in the nose by secreting the serine protease Esp, which can inhibit the adherence proteins secreted by *S. aureus*, inhibiting *S. aureus* biofilm formation and nasal colonization[4]. Another weapon of *S. epidermidis* is the production of antimicrobials for the direct killing of specific species[5]. Moreover, they can also produce quorum-sensing inhibitors to affect the organization of microbiota[6]. Therefore, *S. epidermidis* is actually a bacteria leader dominant in shaping the human microbiota.

The skin, as the largest and exterior organ of humans and the physical barrier towards invasion of foreign pathogens, is home to complex microbial communities. CoNS constitutes a large percentage of the skin microbiota, among which *S. epidermidis* is the most abundant and famous[7]. The surface protein SdrF facilitates the skin colonization of *S. epidermidis*. SdrF is present in 54%–67% of colonizing and clinical isolates[8,9]. SdrF binds human keratins 1 and 10 for adherence to keratinocytes and epithelial cells. SdrF antibodies can reduce the adherence of *S. epidermidis* to keratin and keratinocytes[10]. Almost 100% of the human skin is consistently colonized by *S. epidermidis*[11].

S. epidermidis as pathogen

S. epidermidis does not always behave in the interest of its host. In addition, it is well known that *S. epidermidis* is a leading cause of hospital-acquired infections. *S. epidermidis* causes various diseases, from bloodstream infections to joint infections. Infections caused by *S. epidermidis* are mainly related to the use of medical devices in seriously ill or immunocompromised patients. *S. epidermidis* forms biofilms on these devices, leading to efficient infections. Biofilms are considered the main pathogenic factors of *S. epidermidis*[12]. A biofilm is a bacterial agglomeration within an extracellular matrix that adheres to a surface[13]. This grouped structure protects *S. epidermidis* from antibiotic killing and host immune response and is considered the main mechanism of CoNS infection[13]. Biofilms are formed by the production of poly-N-acetylglucosamine (PNAG) encoded by the *icaADBC* operon[14]. Some *icaADBC*-negative *S. epidermidis* can also form biofilms[15-17]. This is defined as protein-dependent biofilm formation in which the accumulation-associated protein (Aap) plays an important role[18 19]. Proteases can induce biofilm-negative *S. epidermidis* strains to form biofilms[19]. A recent research group

reported that trypsin could significantly increase biofilm formation in commensal *S. epidermidis*, which is biofilm-negative[20]. The whole genome sequences of ATCC 12228 and RP62A, which are a non-biofilm-forming strain and biofilm-forming strain, were published in 2003 and 2005, respectively. The *S. epidermidis* strains ATCC 12228 and RP62A are frequently used as reference strains in various *S. epidermidis* studies

In addition to its biofilm formation ability, another important characteristic of hospital-related *S. epidermidis* is that these infectious strains contain a mobile genetic element, staphylococcal chromosome cassette (SCC)*mec* elements, in their genomes. The SCC*mec* element contains a *mecA* gene encoding methicillin resistance. Moreover, it can also encode a potent peptide toxin called phenol-soluble modulin PSM-*mec*, which mediates neutrophil killing to promote the survival of *S. epidermidis* in the blood, leading to the death of the host[21].

Studies have also found that co-infection of *S. aureus* together with *S. epidermidis* can promote bloodstream infection[22]; the cell wall polymer peptidoglycan from *S. epidermidis* can strongly suppress the production of antimicrobial reactive oxygen species, thus reducing the *S. aureus* infectious dose by over 1,000-fold[22]. *S. epidermidis* is also the most frequent cause of bloodstream infections in infants[13].

In recent years, fast and cost-effective genome sequencing technologies have accelerated research on virulence factors in *S. epidermidis* by analysing the gene levels. The genome composition and functions of *S. epidermidis* were found to vary largely from strain to strain. However, until now, no virulence factors have been clearly discovered that characterize an *S. epidermidis* strain as a pathogenic strain or a commensal strain. Even for the ST2 *S. epidermidis* strains, which are considered to be clinical types frequently causing infections, many were also found colonizing the skin and in the nasal mucosae of healthy individuals[23,24].

S. epidermidis in phage-mediated HGT as a gene reservoir

SCC*mec* with or without the *mecA* gene

Antibiotic resistance genes for several antibiotics are widespread in *S. epidermidis*, including methicillin, oxacilin, rifamycin, fluoroquinolones, gentamycin, tetracycline, chloramphenicol, erythromycin, clindamycin, sulphonamides, and linezolide[25]. According to epidermiology reports worldwide, up to 90% of clinical *S. epidermidis* isolated from hospitals are resistant to methicillin, which is the first line antibiotic against staphylococcal infections. The rate is much higher when compared to that of its more virulent cousin *S. aureus*, which is approximately 50%[26]. Methicillin resistance is mediated by the *mecA* gene, which is located on the mobile genetic element SCC*mec*. The *mecA* gene encodes a penicillin-binding protein, PBP2a, which has a decreased affinity for methicillin compared with the affinities of normal PBPs[27]. It is very likely that *S. aureus* obtained methicillin resistance from *S. epidermidis*. A frequent loss of acquisition of the SCC*mec* element is observed in the *S. epidermidis* population, causing the phenomenon that closely related *S. epidermidis* clones harbour different types of SCC*mec* elements[28]. Until now, ten different SCC*mec* types have been identified in *S. epidermidis*, with type IV being the most common[29]. This type IV SCC*mec* is frequently detected in community-associated methicillin-resistant *S. aureus* (CA-MRSA), which has a greater virulence characteristic compared to hospital-associated methicillin-resistant *S. aureus* (HA-MRSA) and might be an important factor in its widespread in the community[30].

S. aureus ST395 strains present a unique WTA structure that is not in common with other *S. aureus* strains but is similar to CoNS. This CoNS-like WTA structure enables DNA exchange with CoNS via phages[31]. Larsen *et al* found an 88 kb, unusually large SCC*mec* element in one ST395 *S. aureus* strain, JS395, when analysing its genome sequence. A SCC*mec* V element and a CRISPR locus are located 33 kb downstream of *orfX*, which is mostly similar to the SCC*mec* element found in the CoNS species *S. capitis* and *Staphylococcus schleiferi*. Another 55 kb SCC element was identified downstream of this SCC*mec* V element. This SCC element contains a mercury resistance region found in the composite SCC elements of some *S. epidermidis* and *S. aureus* strains. The large size of SCC*mec* suggests that its movement is likely directed by phage-mediated HGT. It is highly possible that *S. aureus* ST395 strains take up DNA from CoNS. These *S. aureus* ST395 strains might serve as an entry point for SCC*mec* and SCC elements from CoNS

into the *S. aureus* population[32].

Linezolid resistance gene

Linezolid is the first agent approved to treat infections caused by MRSA. Linezolid binds to the centre of peptidyl transferase on the ribosome to inhibit the synthesis of bacterial proteins. In recent years, CoNS strains have been isolated from hospitals at an increasing rate due to the frequent use of linezolid[33,34]. The excessive use of linezolid increases the risk of linezolid resistance strains of *S. aureus* and other species, whose resistance rates are now at a low level[35].

Mutations in the 23S rRNA subunit are the most common mechanism of resistance, causing an alteration in the peptidyl transferase centre (PTC), where conserved residues interact directly with linezolid[34]. The acquired resistance mechanism has also been reported recently[36]. Grove *et al.* reported that the acquisition of the natural chloramphenicol-florfenicol resistance (*cfr*) gene carried by a plasmid encodes a protein that catalyses the posttranscriptional methylation of the C-8 atom of A2503, a key residue in 23S rRNA[37]. The *cfr* gene is a highly mobile genetic element among different Staphylococcus species. Until now, the *cfr* gene has been identified in *staphylococci* such as *S. aureus* and CoNS, *enterococci*, and *streptococci*[34]. The methylation by the *cfr* product results in a multidrug resistance phenotype that includes linezolid, lincosamides, and streptogramins[38].

S. epidermidis acts as a reservoir in the spread of linezolid resistance genes. To investigate alternative mechanisms for the spread of the *cfr* gene, Cafini *et al* successfully conducted an *in vitro* horizontal gene transmission of the *cfr* pSCFS7-like plasmid from Spanish *S. epidermidis* ST2 strains to several clinical MRSA strains obtained in Japan under laboratory conditions. They observed that phage-mediated transduction could be an alternative pathway for the dissemination of the *cfr* gene between MRSA strains in addition to the main mechanism of conjugation[39].

SasX

Li *et al* reported the emergence of highly epidemic MRSA ST239 strains carrying a phage harbouring a new cell wall-anchored virulence factor named SasX[40]. SasX promotes nasal colonization, bacterial aggregation, and virulence. The *sasX* gene is located in an unusual 127 kb ΦSPbeta-like prophage with an integrase gene. This prophage is significantly larger than other

typical *S. aureus* siphoviruses. It is highly similar to a prophage located in the genome of *S. epidermidis* RP62A[41]. This is evidence that the ΦSPbeta-like prophage together with the virulence gene *sasX* is acquired from *S. epidermidis*. This phage also spreads between *S. aureus* strains to ST5 lineages[40]. How this *S. epidermidis* phage succeeds in crossing the barrier between different species remains to be investigated. Additionally, the insertion site of this phage within SAR2132 (coding for a putative membrane protein) has thus far not been described for any *S. aureus* phage.

This SasX protein shares high identity with SesI, a virulence determinant harboured in the similar prophage of *S. epidermidis* RP62A. SesI is also a cell wall-associated protein with the leucine–proline–variable amino acid (X)–threonine–glycine (LPXTG) motif also found in SasX[8]. SesI is a marker for invasive *S. epidermidis* isolates since the prevalence of its gene is higher in strains isolated from clinics than in strains isolated from healthy people[42]. In addition, the *S. epidermidis* infection strains, which are *sesI*-positive, also present more virulence genes when compared to *sesI*-negative isolates[43].

Biofilm formation-related genes

S. epidermidis, *S. caprae*, and *S. capitis* are all commensal bacteria on human skin causing opportunistic infections in some circumstances[13,44,45]. Watanabe *et al.* recently compared the genome sequences of *S. caprae* and 24 Staphylococcus species and reported that the genome sequences of *S. caprae* shared high similarities with those of *S. epidermidis* and *S. capitis*[1]. These three human disease-related CoNS species commonly share many biofilm formation-associated genes. These virulence factors are considered the main characteristics of *S. epidermidis*, causing infections associated with indwelling medical devices or catheter use. The factors conserved in the genomes of *S. epidermidis*, *S. caprae* and *S. capitis* include WTA synthesis genes, poly-gamma-DL-glutamic acid capsule synthesis genes, and other genes encoding nonproteinaceous adhesins[1]. Watanabe *et al.* can also detect biofilm formation with these *S. caprae* and *S. capitis* strains under laboratory conditions. The mechanism of HGT for these genes is unknown. From the genome maps shown in the paper, there were Tn554 in *S. epidermidis* RP62A, *S. capitis* TW2795, and *S. caprae* JMUB898.

Phages in bacterial evolution

Bacteriophages (phages) are viruses that infect and reproduce in bacteria. By microbiota analysis, Oh *et al* found that a large number of phages exist in the nose and skin of humans, which are sites that *S. epidermidis* and other Staphylococcal species frequently inhabit[46]. Since phages and *S. epidermidis* share the same niche, we cannot exclude the possibility that phages are involved in the co-evolution of *S. epidermidis* and its relatives in the Staphylococcal genus.

Phage was a hot topic in the scientific research field as early as the 1920s. However, less attention has been paid to *S. epidermidis* phages because *S. epidermidis* was usually considered a commensal bacterium some years before. *S. epidermidis* phages were the first to be used for typing clinical *S. epidermidis* strains. The analysis of the uploaded CoNS genomes found only a few strains containing prophages. However, the prevalence of phages might be underestimated in clinical strains of CoNS[47].

Staphylococcal phages are classified by morphology and serotype. All reported *S. epidermidis* phages belong to the order of *Caudovirales*, composed of an icosahedral capsid filled with double-stranded DNA and a thin filamentous tail. The morphological classification showed that most of the *S. epidermidis* phages belong to the *Siphoviridae* family, composed of an icosahedral capsid and a non-contractile tail that varies from 130-400 nm with a base plate[48]. A small portion of the *S. epidermidis* phage population belongs to *Myoviridae*, which have an icosahedral capsid and a long, contractile, double-sheathed tail, and *Podoviridae*, which have a small icosahedral capsid and a very short, non-contractile tail. Eleven serogroups (A-H and J-L) of Staphylococcal phages have been defined[49,50]. Serogroup E, J and K phages are coagulase-negative staphylococcal phages. For Staphylococcus prophages of the major *Siphoviridae* family, a classification method based on polymorphisms of the integrase gene was developed by Wolz's lab in Tübingen. This method clusters phages with the integrase gene sequences, allowing prediction of the phage insertion sites while also relating the virulence gene content[51].

It is well known that a large number of Staphylococcal strains contain prophages. Phages shape the bacterial genome by moving around bacteria, promoting horizontal gene transfer. Usually, phages carry many virulence or fitness genes. Phages contribute to the induction, packaging, and transfer of genomic islands[52]. Phage transduction is an efficient way to transfer a gene. For *S.*

epidermidis, which is not a naturally competent species, phages are thought to be the major causes of gene evolution and diversification. Phage offer bacteria an opportunity to evolve quickly, helping them adapt to the ever-changing environment. Phages could also disrupt genes in a bacterial genome or protect bacteria from lytic infection by other phages.

Phages provide advantages to bacteria by generating heterogeneity within a bacterial population during the infection process, which helps bacteria adapt efficiently to host defences. Several research papers have shown that the mobilization of phages and genomic integration preferentially occur under infection conditions rather than colonizing conditions in healthy human hosts[51,53,54]. The changing environmental conditions, such as antibiotic treatments or reactive oxygen species released by macrophages, are factors that cause phage induction.

S. epidermidis phages

In recent years, an increasing number of *S. epidermidis* phage genomes have been sequenced and available in GenBank (Table 1). These *S. epidermidis* phages have general genomic features of phages with high gene density and a typical genome organization composed of five functional modules (lysogeny, DNA metabolism, DNA packaging and capsid morphogenesis, tail morphogenesis and lysis)[55]. Comparative analysis of genome sequences revealed that *S. epidermidis* vB_SepiS_phiIPLA5, vB_SepiS_phiIPLA7, PH15, CNPx, CNPH82, IME1348_01, and our ΦTÜB (genome sequence unpublished) are closely related[56,57] (see Chapter 2). Madhusoodanan reported a superantigen-bearing genomic island found in *S. epidermidis*[58]. In addition to staphylococcal enterotoxin C3 (SEC3), *S. epidermidis* also encodes an enterotoxin-like toxin L (SEIL), which is highly similar to the toxin in *S. aureus*. This is evidence that horizontal transfer occurs between different Staphylococcal species.

Phage PH15 is the first *S. epidermidis* phage whose complete genome sequence was analysed and reported[59]. The phylogenetic analysis of all known *Siphoviridae* at that time showed that PH15 was clustered in a novel clade within the *S. aureus* phage group. This indicates a possible genetic exchange between *S. epidermidis* phages and *S. aureus* phages. Recently, we reported a new *S. epidermidis* phage and its ability to transduce plasmids into *S. epidermidis*. This new bacteriophage, named ΦTÜB, was isolated from a clinical *S. epidermidis* strain in Tübingen. Based on electron microscopy, the phage was considered to be a temperate phage of the

Siphoviridae family. Genome sequencing showed that the ΦTÜB genome consists of 44,592 bp of dsDNA with a GC content of 34.5%. Although ΦTÜB showed similarity with the *S. aureus* phage Φ11 in host recognition genes, it can only adsorb to *S. epidermidis* but not to *S. aureus*. It is also interesting that this ΦTÜB has a unique conversion lysogen gene that is lacking in all other *S. epidermidis* phages reported until now.

WTA as a link between *S. epidermidis* and phages

WTA is the most abundant molecule on the cell wall of s*taphylococci* (Figure 1). During adsorption, phages can recognize WTA as their specific receptor, leading to infection and replication. Different Staphylococcal species have different WTA structures (Table 2). All CoNS have similar WTA structures. *S. epidermidis* usually has a WTA structure composed of polymers of 1,3-glycerol-phosphate (GroP) with α/β-glucose or α-N-acetylglucosamine (GlcNAc)[60] (see Chapter 2). *S. capitis* and *S. hominis* have GroP polymers decorated with either α- or β-GlcNAc or both[61]. The WTA of coagulase-positive *S. aureus* is usually composed of a polymer of 1,5-ribitol-phosphate (RboP) decorated with either α- or β-GlcNAc or both[62-64]. Recently, strains of certain sequence types of *S. epidermidis* and *S. aureus* were reported to present special WTA structures. *S. aureus* ST395 presents a WTA structure of GroP-polymer and α-GalNAc similar to CoNS[65]. In *S. epidermidis*, ST23 and some other clones have a *S. aureus*-similar WTA structure of RboP-polymer decorated with α-glucose together with a *S. epidermidis*-common GroP-polymer WTA structure with no sugar decoration (see Chapter 2). Despite the different sugar decorations observed in different Staphylococcal species and strains, all of the WTA polymers reported in s*taphylococci* are modified with D-alanyl.

Table 1. *S. epidermidis* phage genome information from NCBI.

Phage	Family	Genome size(kb)	Protein	Gene	GC content(%)
ΦTÜB	*Siphoviridae*	44.6	71	71	34.5
PH15	*Siphoviridae*	44.0	68	68	34.9
IME1348_01	*Siphoviridae*	42.4	63	63	34.7
CNPH82	*Siphoviridae*	52.9	65	65	34.7
Spbeta-like	*Siphoviridae*	127.7	156	156	30.5
HOB 14.1.R1	*Siphoviridae*	18.7	37	37	32.9
vB_SepS_SEP9	*Siphoviridae*	92.4	129	130	30.5
StB20-like	*Siphoviridae*	40.7	59	59	33.3
IPLA5	*Siphoviridae*	43.6	66	67	34.7
CNPx	*Siphoviridae*	53.3	69	69	34.7
IPLA7	*Siphoviridae*	42.1	59	59	34.8
Terranova	*Herelleviridae*	141.3	205	205	28.0
Twillingate	*Herelleviridae*	142.6	206	207	28.0
Quidividi	*Herelleviridae*	141.5	203	204	28.0
SEP1	*Herelleviridae*	139.9	200	200	27.90
St134	*Podoviridae*	18.3	21	21	30.1
Andhra	*Podoviridae*	18.6	20	20	29.8

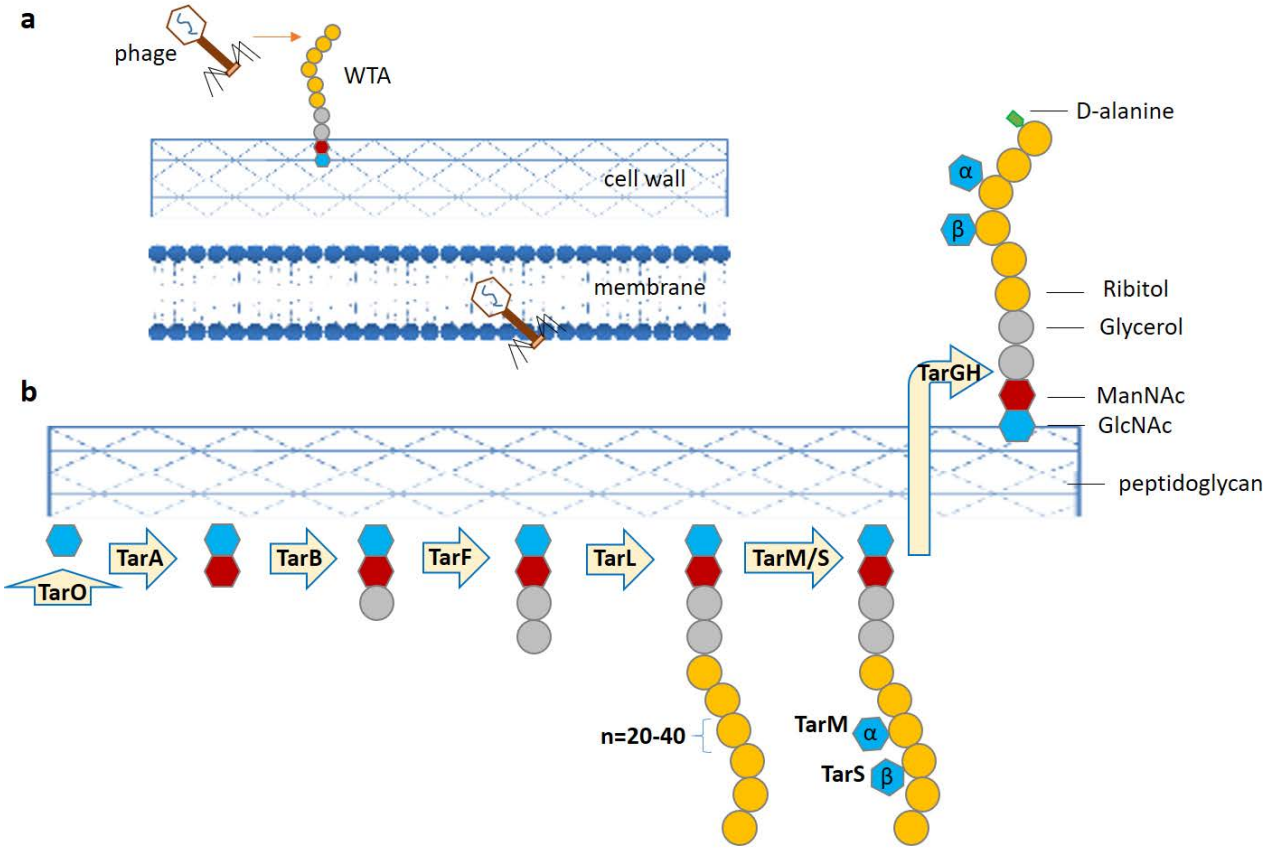


Figure 1. WTAs in gram-positive bacteria. (a) Phage interaction with WTAs. The phage adsorbs to WTAs and injects its DNA inside the bacterial cell. The gram-positive cell wall showing that WTAs are covalently anchored to cell wall peptidoglycan. (b) Schema of the WTA biosynthesis pathway in *S. aureus*.

Table 2. Components of WTA from different staphylococcal species.

Species	Strains	Polymers	Sugar decoration	Reference
S. aureus	RN4220	RboP	α-GlcNAc, β-GlcNAc	Xia *et al*[62]
	NCTC8325	RboP	α-GlcNAc, β-GlcNAc	Jenni *et al*[65]
	PS187	GroP	α-GalNAc	Winstel *et al*[31]
S. epidermidis	RP62A	GroP	α-Glc,α-GlcNAc	Sadovskaya *et al*[59]
	1457	GroP	α-Glc	this thesis (Chapter 2)
	E73	GroP, RboP	α-Glc	this thesis (Chapter 2)
S. capitis	ATCC 27840	GroP	α-GlcNAc	Endl *et al*[60]
	Kloos LK242	GroP	α-GlcNAc, β-GlcNAc	Endl *et al*
S. haemolyticus	DSM 20264	GroP	α-GlcNAc, β-GlcNAc	Endl *et al*
	Kloos WK246	GroP	α-GlcNAc, β-GlcNAc	Endl *et al*
S. hominis	ATCC 27844	GroP	α-GlcNAc, β-GlcNAc	Endl *et al*
	Kloos MK129	GroP	α-GlcNAc	Endl *et al*
S. carnosus	TM300	GroP	Glc, GalNAc	Schleifer *et al*[66]
S. lugdunensis	N920243	GroP	(GalNAc)	Mnich *et al*[67]
S. cohnii	DSM 20260	GroP	α-Glc,α-GlcNAc	Endl *et al*
	DM224		β-Glc,α-GlcNAc	Endl *et al*
S. waneri	Kloos RM130	GroP	α-Glc,β-GlcNAc	Endl *et al*
	ATCC 27836	GroP	β-Glc,β-GlcNAc	Endl *et al*
S. intermedius	Hajek K6	GroP	α-Glc,α-GlcNAc	Endl *et al*
	Hajek K2	GroP	α-GlcNAc, β-GlcNAc	Endl *et al*
S. simulans	ATCC 27848	GroP	α-GalNAc, α,β-GlcNAc	Endl *et al*
	Kloos KL299	GroP	α-GalNAc, β-GlcNAc	Endl *et al*
S. saprophycus	Kloos KL251	GroP	α-GlcNAc, β-GlcNAc	Endl *et al*
	Kloos KL122	GroP	α-GlcNAc, β-GlcNAc	Endl *et al*
		RboP	β-GlcNAc	
S. xylosus	DSM20267	GroP	α-GlcNAc, β-GlcNAc	Endl *et al*
		RboP	β-GlcNAc	
	CCM 1400	GroP	α-GlcNAc	Endl *et al*
		RboP	β-GlcNAc	

Aim of this thesis

S. epidermidis is both an important commensal bacterium in humans and a pathogen as one of the leading causes of hospital infections. Whether certain kinds of *S. epidermidis* are more virulent and the underlying mechanisms is a hot topic in the scientific field of Staphylococcal study. An accurate diagnostic method to distinguish true pathogens from sample contaminants of colonization strains of *S. epidermidis* is needed; therefore, the identification of a biomarker is required. None of the virulence factors discovered until now could categorize a certain *S. epidermidis* strain as pathogenic or commensal. Although *S. epidermidis* ST2 and biofilm formation are frequently found in the hospital environment, the direct correlation between the genes involved in the virulence of *S. epidermidis* has never been confirmed.

The development of sequence technologies provides us with more genome sequences of *S. epidermidis*. *S. epidermidis* is considered a gene reservoir for *S. aureus* and other s*taphylococci*. More and more similar genetic elements have been found in *S. epidermidis* and its cousin *S. aureus* or other s*taphylococci*. The real fact and mechanism of how genes transfer between *S. epidermidis* and other s*taphylococci* is still unknown. The SCC*mec* element, as one of the most important genetic elements shared, is large (20-70 kb) and can only be moved by phages. However, this seems to be not possible because the specific phage receptors on *S. epidermidis* and *S. aureus* were reported to be different types, poly(GroP) WTA and poly(RboP) WTA, respectively.

To answer these questions, this thesis followed the major aims listed here to study the newly discovered *S. epidermidis* strains with two types of WTA, expressing a second *S. aureus*-like poly(RboP) WTA (main research procedures shown in Figure 3):

(i) Elucidate the role of RboP WTA in the colonization and infection capacity of *S. epidermidis*. We approached this question by constructing a *S. aureus*-like poly(RboP) WTA mutant with one of the two types of WTA from the *S. epidermidis* strain E73, a parental complement strain with pRB474-*tarIJLM2* that is the related genes of poly(RboP) WTA, and a 1457 poly(RboP) WTA expressing strain with pRB474-*tarIJLM2*. Human epithelial cells and cotton rat nasal cells were utilized for the *in vitro* binding model. *In vivo* models were also studied with both cotton rats and Balb/C mice. Biofilm formation and oxacillin resistance were also tested with these two WTA

strains.

(ii) Elucidate the prevalence and mobility of the RboP WTA genes in *S. epidermidis* clonal lineages. For the purpose of checking the biomarker potential of the related *tarIJLM2* gene cluster and its prevalence, both clinical *S. epidermidis* infection strains and colonization strains from healthy volunteers were collected from both Germany and China. Sequence types of these strains were analysed. A large number of online genome data were also analysed. Bioinformatic methods were used to study the genomic environments of *tarIJLM2* to determine the possible origin and mobile potential of this new gene cluster.

(iii) Elucidate the advantage of *S. epidermidis* producing two types of WTA in promoting interspecies horizontal gene transfer of s*taphylococci*. A SaPI assay using *S. aureus* phage Φ11 as a helper phage to transduce SaPI to these *S. epidermidis* and their WTA mutants was performed. Φ187, which is a GroP-specific phage, was also used in the SaPI assay to mimic gene transduction from *S. aureus* ST395 to these *S. epidermidis* strains with two types of WTA. We also attempted to find novel *S. epidermidis* phages that could transduce and spread the SaPI from these *S. epidermidis* strains with two types of WTA to other *S. epidermidis*.

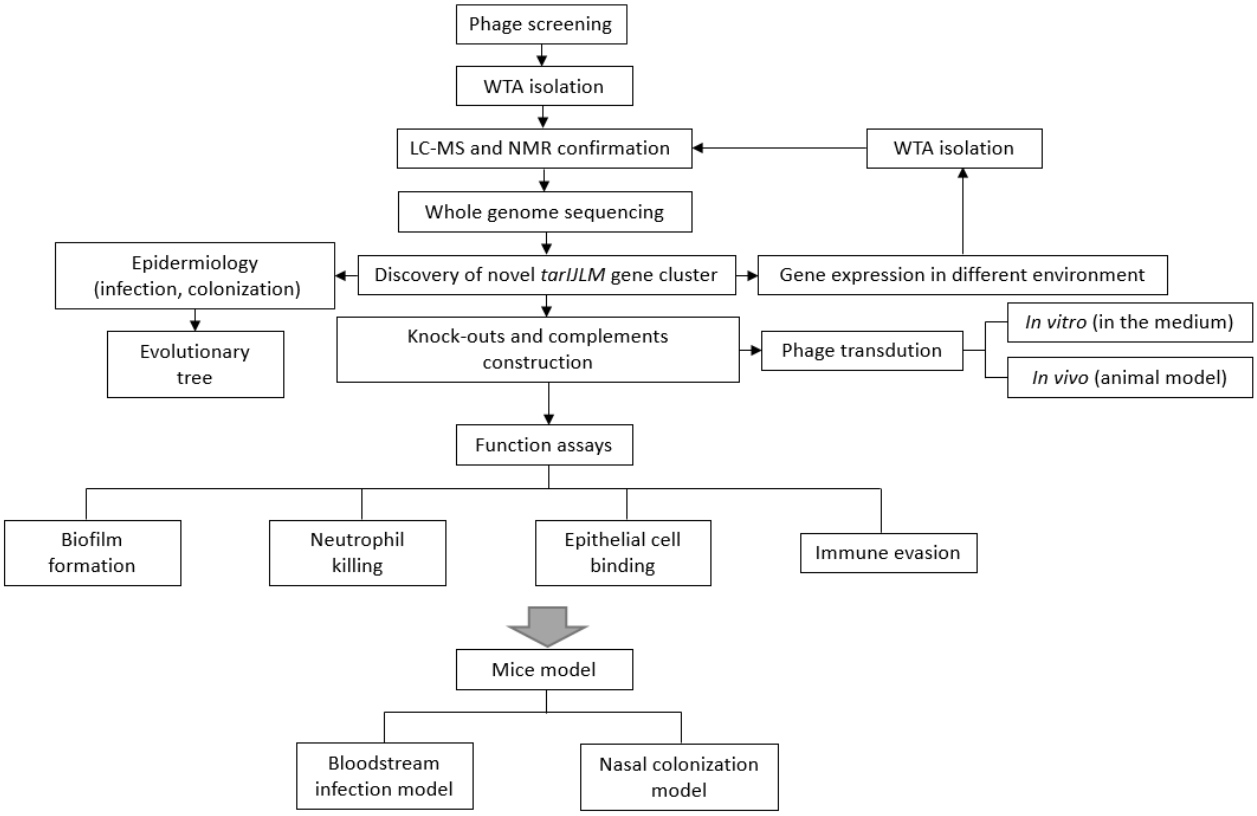


Figure 2. Main working flow of the main project (chapter 2) of my Ph.D. thesis.

References

1 Watanabe, S. *et al.* Complete genome sequencing of three human clinical isolates of Staphylococcus caprae reveals virulence factors similar to those of S-epidermidis and S-capitis. *Bmc Genomics* **19**, doi:ARTN 810.10.1186/s12864-018-5185-9 (2018).

2 Liu, C. M. *et al.* Staphylococcus aureus and the ecology of the nasal microbiome. *Sci Adv* **1**, doi:ARTN e1400216. 10.1126/sciadv.1400216 (2015).

3 Sakr, A. *et al.* Staphylococcus aureus Nasal Colonization: An update on Mechanisms, Epidemiology, Risk Factors, and Subsequent Infections. Front. Microbiol. 9:2419. doi:10.3389/fmicb.2018.02419 (2018).

4 Iwase, T. *et al.* Staphylococcus epidermidis Esp inhibits Staphylococcus aureus biofilm formation and nasal colonization. *Nature* **465**, 346-349, doi:10.1038/nature09074 (2010).

5 Krismer, B., Weidenmaier, C., Zipperer, A. & Peschel, A. The commensal lifestyle of Staphylococcus aureus and its interactions with the nasal microbiota. *Nature Reviews Microbiology* **15**, 675-687, doi:10.1038/nrmicro.2017.104 (2017).

6 Paharik, A. E. *et al.* Coagulase-Negative Staphylococcal Strain Prevents Staphylococcus aureus Colonization and Skin Infection by Blocking Quorum Sensing. *Cell Host Microbe* **22**, 746-+, doi:10.1016/j.chom.2017.11.001 (2017).

7 Schleifer, K. H. & Kloos, W. E. Isolation and Characterization of Staphylococci from Human Skin .1. Amended Descriptions of Staphylococcus-Epidermidis and Staphylococcus-Saprophyticus and Descriptions of 3 New Species - Staphylococcus-Cohnii, Staphylococcus-Haemolyticus, and Staphylococcus-Xylosus. *Int J Syst Bacteriol* **25**, 50-61, doi:Doi 10.1099/00207713-25-1-50 (1975).

8 Bowden, M. G. *et al.* Identification and preliminary characterization of cell-wall-anchored proteins of Staphylococcus epidermidis. *Microbiol-Sgm* **151**, 1453-1464, doi:10.1099/mic.0.27534-0 (2005).

9 Arrecubieta, C. *et al.* SdrF, a Staphylococcus epidermidis Surface Protein, Contributes to the Initiation of Ventricular Assist Device Driveline-Related Infections. *Plos Pathogens* **5**, doi:ARTN e1000411. 10.1371/journal.ppat.1000411 (2009).

10 Trivedi, S. *et al.* The Surface Protein SdrF Mediates Staphylococcus epidermidis Adherence to Keratin. *Journal of Infectious Diseases* **215**, 1846-1854, doi:10.1093/infdis/jix213 (2017).

11 Grice, E. A. & Segre, J. A. The skin microbiome (vol 9, pg 244, 2011). *Nature Reviews Microbiology* **9**, doi:10.1038/nrmicro2619 (2011).

12 Le, K. Y., Park, M. D. & Otto, M. Immune Evasion Mechanisms of Staphylococcus epidermidis Biofilm Infection. *Frontiers in Microbiology* **9**, doi:ARTN 359.10.3389/fmicb.2018.00359 (2018).

13 Otto, M. Staphylococcus epidermidis--the 'accidental' pathogen. *Nat Rev Microbiol* **7**, 555-567, doi:10.1038/nrmicro2182 (2009).

14 Ziebuhr, W. *et al.* Detection of the intercellular adhesion gene cluster (ica) and phase variation in Staphylococcus epidermidis blood culture strains and mucosal isolates. *Infect Immun* **65**, 890-896 (1997).

15 Hellmark, B., Soderquist, B., Unemo, M. & Nilsdotter-Augustinsson, A. Comparison of Staphylococcus epidermidis isolated from prosthetic joint infections and commensal isolates in regard to antibiotic susceptibility, agr type, biofilm production, and epidemiology. *International Journal of Medical Microbiology* **303**, 32-39, doi:10.1016/j.ijmm.2012.11.001 (2013).

16 Dice, B. *et al.* Biofilm formation by ica-positive and ica-negative strains of Staphylococcus epidermidis in vitro. *Biofouling* **25**, 367-375, doi:10.1080/08927010902803297 (2009).

17 Rohde, H. *et al.* Polysaccharide intercellular adhesin or protein factors in biofilm accumulation of Staphylococcus epidermidis and Staphylococcus aureus isolated from prosthetic hip and knee joint infections. *Biomaterials* **28**, 1711-1720, doi:10.1016/j.biomaterials.2006.11.046 (2007).

18 Hennig, S., Wai, S. N. & Ziebuhr, W. Spontaneous switch to PIA-independent biofilm formation in an ica-positive Staphylococcus epidermidis isolate. *International Journal of Medical Microbiology* **297**, 117-122, doi:10.1016/j.ijmm.2006.12.001 (2007).

19 Rohde, H. *et al.* Detection of virulence-associated genes not useful for discriminating between invasive and commensal Staphylococcus epidermidis strains from a bone marrow transplant unit. *J Clin Microbiol* **42**, 5614-5619, doi:10.1128/JCM.42.12.5614-5619.2004 (2004).

20 Martinez-Garcia, S. *et al.* Non-biofilm-forming commensal Staphylococcus epidermidis isolates produce biofilm in the presence of trypsin. *Microbiologyopen*, doi:ARTN e906.10.1002/mbo3.906 (2019).

21 Qin, L. *et al.* Toxin Mediates Sepsis Caused by Methicillin-Resistant Staphylococcus epidermidis. *Plos Pathogens* **13**, doi:ARTN e1006153.10.1371/journal.ppat.1006153 (2017).

22 Boldock, E. *et al.* Human skin commensals augment Staphylococcus aureus pathogenesis. *Nature Microbiology* **3**, 881-890, doi:10.1038/s41564-018-0198-3 (2018).

23 Miragaia, M., Thomas, J. C., Couto, I., Enright, M. C. & de Lencastre, H. Inferring a population structure for Staphylococcus epidermidis from multilocus sequence typing data. *J Bacteriol* **189**, 2540-2552, doi:10.1128/JB.01484-06 (2007).

24 Flores-Paez, L. A. *et al.* Molecular and Phenotypic Characterization of Staphylococcus epidermidis Isolates from Healthy Conjunctiva and a Comparative Analysis with Isolates from Ocular Infection. *PLoS One* **10**, e0135964, doi:10.1371/journal.pone.0135964 (2015).

25 Rogers, K. L., Fey, P. D. & Rupp, M. E. Coagulase-negative staphylococcal infections. *Infect Dis Clin North Am* **23**, 73-98, doi:10.1016/j.idc.2008.10.001 (2009).

26 Diekema, D. J. *et al.* Survey of infections due to Staphylococcus species: frequency of occurrence and antimicrobial susceptibility of isolates collected in the United States,

Canada, Latin America, Europe, and the Western Pacific region for the SENTRY Antimicrobial Surveillance Program, 1997-1999. *Clin Infect Dis* **32 Suppl 2**, S114-132, doi:10.1086/320184 (2001).

27 Chambers, H. F., Hartman, B. J. & Tomasz, A. Increased amounts of a novel penicillin-binding protein in a strain of methicillin-resistant Staphylococcus aureus exposed to nafcillin. *J Clin Invest* **76**, 325-331, doi:10.1172/JCI111965 (1985).

28 Miragaia, M., Couto, I. & de Lencastre, H. Genetic diversity among methicillin-resistant Staphylococcus epidermidis (MRSE). *Microb Drug Resist* **11**, 83-93, doi:10.1089/mdr.2005.11.83 (2005).

29 International Working Group on the Classification of Staphylococcal Cassette Chromosome, E. Classification of staphylococcal cassette chromosome mec (SCCmec): guidelines for reporting novel SCCmec elements. *Antimicrob Agents Chemother* **53**, 4961-4967, doi:10.1128/AAC.00579-09 (2009).

30 Ma, X. X. *et al.* Novel type of staphylococcal cassette chromosome mec identified in community-acquired methicillin-resistant Staphylococcus aureus strains. *Antimicrob Agents Chemother* **46**, 1147-1152, doi:10.1128/aac.46.4.1147-1152.2002 (2002).

31 Winstel, V. *et al.* Wall Teichoic Acid Glycosylation Governs Staphylococcus aureus Nasal Colonization. *MBio* **6**, e00632, doi:10.1128/mBio.00632-15 (2015).

32 Larsen, J., Andersen, P. S., Winstel, V. & Peschel, A. Staphylococcus aureus CC395 harbours a novel composite staphylococcal cassette chromosome mec element. *J Antimicrob Chemother* **72**, 1002-1005, doi:10.1093/jac/dkw544 (2017).

33 Song, Y. *et al.* cfr-mediated linezolid-resistant clinical isolates of methicillin-resistant coagulase-negative staphylococci from China. *J Glob Antimicrob Resist* **8**, 1-5, doi:10.1016/j.jgar.2016.09.008 (2017).

34 Mendes, R. E., Deshpande, L. M. & Jones, R. N. Linezolid update: stable in vitro activity following more than a decade of clinical use and summary of associated resistance mechanisms. *Drug Resist Updat* **17**, 1-12, doi:10.1016/j.drup.2014.04.002 (2014).

35 Monaco, M., Pimentel de Araujo, F., Cruciani, M., Coccia, E. M. & Pantosti, A. Worldwide Epidemiology and Antibiotic Resistance of Staphylococcus aureus. *Curr Top Microbiol Immunol* **409**, 21-56, doi:10.1007/82_2016_3 (2017).

36 Arias, C. A. *et al.* Clinical and microbiological aspects of linezolid resistance mediated by the cfr gene encoding a 23S rRNA methyltransferase. *J Clin Microbiol* **46**, 892-896, doi:10.1128/JCM.01886-07 (2008).

37 Grove, T. L. *et al.* A substrate radical intermediate in catalysis by the antibiotic resistance protein Cfr. *Nat Chem Biol* **9**, 422-427, doi:10.1038/nchembio.1251 (2013).

38 Toh, S. M. *et al.* Acquisition of a natural resistance gene renders a clinical strain of methicillin-resistant Staphylococcus aureus resistant to the synthetic antibiotic linezolid. *Mol Microbiol* **64**, 1506-1514, doi:10.1111/j.1365-2958.2007.05744.x (2007).

39 Cafini, F. *et al.* Horizontal gene transmission of the cfr gene to MRSA and Enterococcus: role of Staphylococcus epidermidis as a reservoir and alternative pathway for the spread

of linezolid resistance. *J Antimicrob Chemother* **71**, 587-592, doi:10.1093/jac/dkv391 (2016).

40 Li, M. *et al.* MRSA epidemic linked to a quickly spreading colonization and virulence determinant. *Nature Medicine* **18**, 816-U217, doi:10.1038/nm.2692 (2012).

41 Holden, M. T. G. *et al.* Genome Sequence of a Recently Emerged, Highly Transmissible, Multi-Antibiotic- and Antiseptic-Resistant Variant of Methicillin-Resistant Staphylococcus aureus, Sequence Type 239 (TW). *Journal of Bacteriology* **192**, 888-892, doi:10.1128/Jb.01255-09 (2010).

42 Soderquist, B. *et al.* Staphylococcus epidermidis surface protein I (SesI): a marker of the invasive capacity of S. epidermidis? *Journal of Medical Microbiology* **58**, 1395-1397, doi:10.1099/jmm.0.008771-0 (2009).

43 Qi, X. Q. *et al.* SesI May Be Associated with the Invasiveness of Staphylococcus epidermidis. *Frontiers in Microbiology* **8**, doi:ARTN 2574.10.3389/fmicb.2017.02574 (2018).

44 Seng, P. *et al.* Staphylococcus caprae bone and joint infections: a re-emerging infection? *Clin Microbiol Infec* **20**, O1052-O1058, doi:10.1111/1469-0691.12743 (2014).

45 Tevell, S., Hellmark, B., Nilsdotter-Augustinsson, A. & Soderquist, B. Staphylococcus capitis isolated from prosthetic joint infections. *Eur J Clin Microbiol* **36**, 115-122, doi:10.1007/s10096-016-2777-7 (2017).

46 Oh, J. *et al.* Biogeography and individuality shape function in the human skin metagenome. *Nature* **514**, 59-64, doi:10.1038/nature13786 (2014).

47 Lina, B. *et al.* Role of Bacteriophages in Genomic Variability of Related Coagulase-Negative Staphylococci. *Fems Microbiol Lett* **109**, 273-277, doi:DOI 10.1111/j.1574-6968.1993.tb06180.x (1993).

48 Ackermann, H. W. Tailed bacteriophages: The order Caudovirales. *Adv Virus Res* **51**, 135-201 (1999).

49 Rippon, J. E. A new serological division of Staphylococcus aureus bacteriophages: group G. *Nature* **170**, 287, doi:10.1038/170287a0 (1952).

50 Rippon, J. E. The classification of bacteriophages lysing staphylococci. *J Hyg (Lond)* **54**, 213-226, doi:10.1017/s0022172400044478 (1956).

51 Goerke, C. *et al.* Diversity of prophages in dominant Staphylococcus aureus clonal lineages. *J Bacteriol* **191**, 3462-3468, doi:10.1128/JB.01804-08 (2009).

52 Christie, G. E. & Dokland, T. Pirates of the Caudovirales. *Virology* **434**, 210-221, doi:10.1016/j.virol.2012.10.028 (2012).

53 Goerke, C., Wirtz, C., Fluckiger, U. & Wolz, C. Extensive phage dynamics in Staphylococcus aureus contributes to adaptation to the human host during infection. *Mol Microbiol* **61**, 1673-1685, doi:10.1111/j.1365-2958.2006.05354.x (2006).

54 Goerke, C. & Wolz, C. Regulatory and genomic plasticity of Staphylococcus aureus during persistent colonization and infection. *Int J Med Microbiol* **294**, 195-202, doi:10.1016/j.ijmm.2004.06.013 (2004).

55 Deghorain, M. & Van Melderen, L. The Staphylococci phages family: an overview. *Viruses* **4**, 3316-3335 (2012).

56 Rosenstein, R. *et al.* Genome analysis of the meat starter culture bacterium Staphylococcus carnosus TM300. *Appl Environ Microbiol* **75**, 811-822, doi:10.1128/AEM.01982-08 (2009).

57 Gutierrez, D., Martinez, B., Rodriguez, A. & Garcia, P. Genomic characterization of two Staphylococcus epidermidis bacteriophages with anti-biofilm potential. *BMC Genomics* **13**, 228, doi:10.1186/1471-2164-13-228 (2012).

58 Madhusoodanan, J. *et al.* An Enterotoxin-Bearing Pathogenicity Island in Staphylococcus epidermidis. *J Bacteriol* **193**, 1854-1862, doi:10.1128/JB.00162-10 (2011).

59 Daniel, A., Bonnen, P. E. & Fischetti, V. A. First complete genome sequence of two Staphylococcus epidermidis bacteriophages. *J Bacteriol* **189**, 2086-2100, doi:10.1128/JB.01637-06 (2007).

60 Sadovskaya, I., Vinogradov, E., Li, J. J. & Jabbouri, S. Structural elucidation of the extracellular and cell-wall teichoic acids of Staphylococcus epidermidis RP62A, a reference biofilm-positive strain. *Carbohyd Res* **339**, 1467-1473, doi:10.1016/j.carres.2004.03.017 (2004).

61 Endl, J., Seidl, P. H., Fiedler, F. & Schleifer, K. H. Determination of Cell-Wall Teichoic-Acid Structure of Staphylococci by Rapid Chemical and Serological Screening Methods. *Arch Microbiol* **137**, 272-280, doi:Doi 10.1007/Bf00414557 (1984).

62 Weidenmaier, C. & Peschel, A. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nature Reviews Microbiology* **6**, 276-287, doi:10.1038/nrmicro1861 (2008).

63 Xia, G. *et al.* Glycosylation of wall teichoic acid in Staphylococcus aureus by TarM. *J Biol Chem* **285**, 13405-13415, doi:10.1074/jbc.M109.096172 (2010).

64 Brown, S. *et al.* Methicillin resistance in Staphylococcus aureus requires glycosylated wall teichoic acids. *Proc Natl Acad Sci U S A* **109**, 18909-18914, doi:10.1073/pnas.1209126109 (2012).

65 Winstel, V. *et al.* Wall teichoic acid structure governs horizontal gene transfer between major bacterial pathogens. *Nat Commun* **4**, 2345, doi:10.1038/ncomms3345 (2013).

Chapter 2
-
Nosocomial *Staphylococcus epidermidis* remodels surface glycopolymers to shift from commensal to pathogen behavior

Xin Du[1,2], Jesper Larsen[3], Min Li [4], Axel Walter[1,2], Anna Both[5], Patricia M. Sanchez Carballo[6], Marc Stegger[3], Yao Liu[4], Junlan Liu[4], Jessica Schade[7], Katarzyna A.Duda[8], Bernhard Krismer[1,2], Simon Heilbronner[1,2], Christopher Weidenmaier[7], Christoph Mayer[1,2], Holger Rohde[5], Volker Winstel[1,2,9], Andreas Peschel[1,2]

1. Interfaculty Institute of Microbiology and Infection Medicine, Infection Biology, University of Tübingen, Germany. 2. German Center for Infection Research (DZIF), partner site Tübingen, Germany. 3. Statens Serum Institut, Denmark. 4. Renji Hospital, Medical of School, Shanghai Jiaotong University, China. 5. Medical Microbiology Institute, Hospital of Hamburg-Eppendorf University, Germany. 6. Research Group Clinical Infectious Diseases, Research Center Borstel, Leibniz Lung Center, German Center for Infection Research (DZIF), partner site Borstel, Germany. 7. Interfaculty Institute of Microbiology and Infection Medicine, Medical Microbiology, University of Tübingen, Germany. 8. Junior Research Group of Allergobiochemistry, Research Center Borstel, Leibniz Lung Center, Airway Research North (ARCN), German Center for Lung Research (DZL), Borstel, Germany. 9. Current address: Department of Microbiology, University of Chicago, Chicago, Illinois USA.

Ready to be submitted. Target journal: *Nature Medicine*.

Introductory paragraph

Globally spreading healthcare-associated and methicillin-resistant *Staphylococcus epidermidis* (MRSE) clones are major causes of difficult-to-treat implant and bloodstream infections[1,2]. Specific virulence factors distinguishing these from commensal skin and nose-colonizing *S. epidermidis* clones have remained unknown[3,4]. Poly-glycerolphosphate (GroP) wall teichoic acid (WTA) polymers cover the surface of *S. epidermidis*, potentially shaping critical host interactions[5]. We demonstrate that only nosocomial *S. epidermidis* contain an accessory genetic element, *tarIJLM*, to express a second WTA type (poly-ribitolphosphate, RboP) that impairs nasal colonization but augments endothelial attachment and host mortality in experimental bloodstream infections. Moreover, RboP-WTA enables *S. epidermidis* to exchange DNA with *Staphylococcus aureus* via bacteriophages thereby creating routes for the inter-species exchange of methicillin resistance, virulence, and colonization factors[6]. Hence, *tarIJLM* represents a promising biomarker for assessing the pathogenicity and monitoring the evolution of epidemic staphylococcal pathogens and RboP-WTA may become an attractive target for preventive and therapeutic interventions against MRSE infections.

Main text

All known *Staphylococcus* bacteriophages use the species- or strain-specific structure of WTA to detect and bind to appropriate bacterial host cells[7]. Phages are also the major vehicles for horizontal gene transfer (HGT) in s*taphylococci*, governing the evolution of new clonal lineages[8], for instance by exchanging staphylococcal cassette chromosome *mec* (SCC*mec*) elements conferring methicillin resistance[9], *S. aureus* pathogenicity islands (SaPIs)[10], or accessory adhesion factor genes such as *sasX*[11] between *S. aureus* and *S. epidermidis* or other coagulase-negative s*taphylococci* (CoNS). However, it has remained enigmatic how such transducing

phages can traverse the WTA-specific HGT boundaries.

We have previously shown that certain WTA-biosynthetic genes reside on genomic elements that seem to move between different *Staphylococcus* species resulting, for instance, in *S. aureus* lineages with *tagNDF* genes that produce GroP-WTA and share susceptibility to specific transducing phages with GroP-WTA-containing CoNS[12]. Along this line, we hypothesized that some *S. epidermidis* strains may have gained the capacity to exchange DNA via typical *S. aureus* phages. A collection of *S. epidermidis* isolates was screened for their capacity to take up a SaPI via the typical *S. aureus* siphophage Φ11[12]. The vast majority of the isolates could not be SaPI-transduced by Φ11, which is in agreement with the absence of the RboP-WTA receptor structure at the *S. epidermidis* surface. However, some isolates could be transduced (Table 1, Fig. 1a,). These isolates also bound Φ11 (Extended Data Fig. 1), suggesting that they express the phage receptor.

The genome of one of the Φ11-transducible isolates, E73, was sequenced and found to belong to the healthcare-associated MRSE multi-locus sequence type (ST) 23[1,2]. It contained the genes for synthesis of the WTA linkage unit (*tagOADB*) and the GroP-WTA polymerase gene *tagF*[13], present in virtually all *S. epidermidis* genomes. Surprisingly, all available *S. epidermidis* genomes also contained a gene cluster (*tarIJL1*; Fig 1b) encoding homologs of *S. aureus* proteins responsible for RboP synthesis (TarI, TarJ) and polymerization (TarL)[13]. However, expression of *tarL1* was very low or undetectable in the *S. epidermidis* laboratory strains RP62A, 1457, and ATCC12228, or in isolates E6, E45, and E73 cultivated in broth, synthetic nasal medium[14], or human serum (Fig. 1c; Extended Data Fig. 2), suggesting that *tarIJL1* is a pseudogene cluster. Notably, E73 contained a second putative RboP-WTA gene cluster (*tarIJLM2*), encoding also a

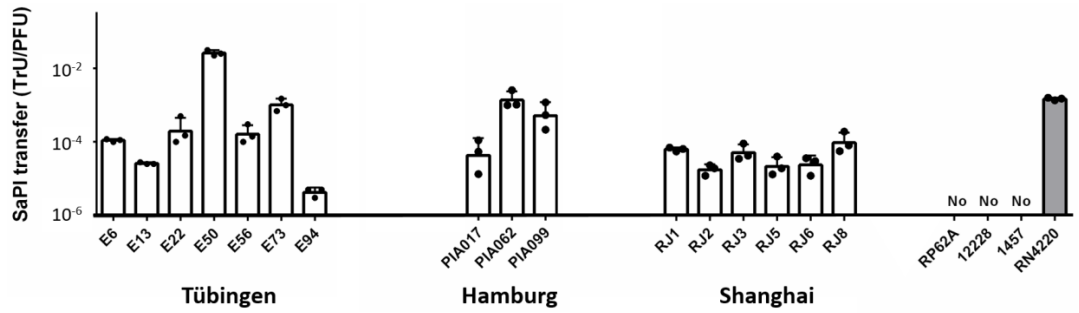
homolog of the *S. aureus* RboP-WTA glycosyl transferase TarM[15]. The *tarIJLM2* cluster replaced a chromosomal gene of unknown function found in most other *S. epidermidis* genomes (Fig. 1b, Supplementary Results). *tarL2* was efficiently expressed in isolates E6, E45, and E73 at all tested growth conditions (Fig. 1c, Extended Data Fig. 2). All but one Φ11-transducible isolates contained *tarIJLM*2 (Table 1), indicating that the rare capacity to acquire foreign DNA from *S. aureus* may depend on the presence of *tarIJLM*2.

Table. 1. Presence of *tarL2* and susceptibility to Φ11-mediated SaPIbov1 transduction in invasive isolates from *S. epidermidis* strain collections.

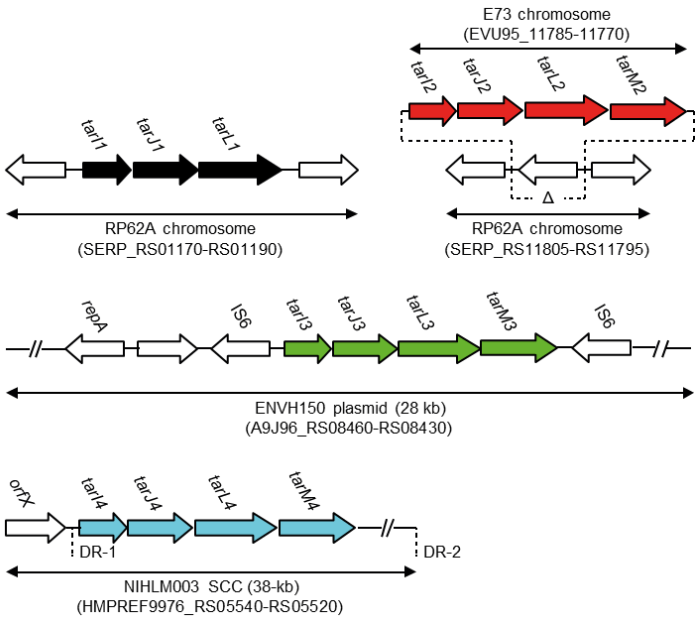
Strain collection	Number of strains			*tarL2*-positive sequence types (numbers)
	Total	*tarL2*-positive	SaPIbov1-transduction by Φ11	
Tübingen				ST5 (2), ST10 (1), ST23 (4),
Nose	155	0	0	ST87 (3)
Skin	144	0	0	
Infection	72	10 (13.9%)	7 (9.7%)[#]	
Hamburg				ST10 (1), ST23 (3), ST87 (2)
Nose	37	0	0	
Infection	75	6 (8.0%)	3 (4.0%)[#]	
Shanghai				ST2 (5)*, single-locus ST23 variant (3)
Nose	206	0	0	
Skin	138	0	0	
Infection	130	8 (6.2%)*	6 (4.6%)[#]	
Total	957	24 (2.5%)	16 (1.7%)[#]	

[#], it remains unclear why some *tarIJLM*2-positive isolates could not be transduced by Φ11, although they all bound Φ11 (Extended Data Fig. 1a). Potential reasons may include specific restriction barrier or clustered regularly interspaced short palindromic repeats (CRISPR) Cas systems.*, one ST2 isolate included was positive for *tarIJLM4* and Φ11 transduction.

a


SaPI transfer (TrU/PFU)
10-2
10-4
10-6
E6
E13
E22
E50
E56
E73
E94
PIA017
PIA062
PIA099
RJ1
RJ2
RJ3
RJ5
RJ6
RJ8
No
No
No
RP62A
12228
1457
RN4220
Tübingen
Hamburg
Shanghai

b


E73 chromosome
(EVU95_11785-11770)
tarI2
tarJ2
tarL2
tarM2
tarI1
tarJ1
tarL1
Δ
RP62A chromosome
(SERP_RS01170-RS01190)
RP62A chromosome
(SERP_RS11805-RS11795)
repA
IS6
tarI3
tarJ3
tarL3
tarM3
IS6
ENVH150 plasmid (28 kb)
(A9J96_RS08460-RS08430)
orfX
tarI4
tarJ4
tarL4
tarM4
DR-1
DR-2
NIHLM003 SCC (38-kb)
(HMPREF9976_RS05540-RS05520)

c

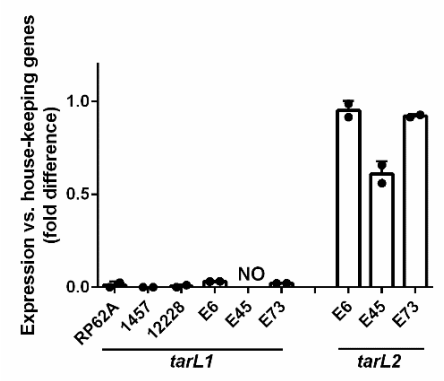

Expression vs. house-keeping genes
(fold difference)
1.0
0.5
0.0
NO
RP62A
1457
12228
E6
E45
E73
E6
E45
E73
tarL1
tarL2

Fig. 1. Some *S. epidermidis* genomes encode *tarIJLM* gene clusters permitting transduction by *S. aureus* phage Φ11. (a) Some infection-derived *S. epidermidis* isolates from Tübingen, Hamburg, and Shanghai can be transduced with SaPIbov by Φ11. Values indicate the ratio of transduction units (TrU) to plaque-forming units (PFU) given as means ± s.d. of three independent experiments. No, no transductants obtained. The origin and properties of the isolates are described in Extended Data Table 1. RP62A, ATCC12228, and 1457 are *S. epidermidis* laboratory strains, only expressing GroP-WTA. *S. aureus* RN4220 expresses RboP-WTA. (b) Potential RboP-WTA biosynthetic gene clusters found in the genomes of all *S. epidermidis* (*tarIJL1*) or certain STs (*tarIJLM2-4*). The genetic location is shown in representative strains RP62A (ST10), E73 (ST23), ENVH150 (ST2), and NIHLM003 (ST218). The direct repeat sequences (DR-1 and DR-2) flanking the SCC element in NIHLM003 are indicated. The genomic location of *tarIJLM5* genes in isolate SNUC_2569 could not be determined due to their presence on two short contigs with no flanking sequences. The figure is not drawn to scale. (c) While *tarL2* is efficiently transcribed in the *tarIJLM2*-positive isolates E6, E45, and E73, *tarIJL1* is not or only very weakly expressed in all tested *S. epidermidis* strains during growth in TSB after 6 h of growth. Values represent means ± s.d. of two independent experiments. They were normalized for strongly and constitutively expressed house-keeping genes *gyrB*, *rho*, and *tpiA*. NO: below detection limit.

WTA isolated from the *tarIJLM*2-negative *S. epidermidis* laboratory strains RP62A, ATCC12228, and 1457 was found to contain only GroP-WTA, as expected, while that of E73 contained two types of WTA, GroP-WTA and RboP-WTA (Fig. 2a, Extended Data Fig. 3). Strain 1457 transformed with a *tarIJLM*2-expressing plasmid produced both, GroP- and RboP-WTA (Fig. 2b, Extended Data Fig. 3) and could be transduced by Φ11 (Fig. 2c) thereby confirming that *tarIJLM2* confers the capacity to produce RboP-WTA and enables *S. epidermidis* to exchange DNA with *S. aureus* via RboP-WTA-binding phages. Moreover, deletion of *tarIJLM*2 in E73 abolished RboP-WTA production and transduction as well as binding by Φ11, which was restored by complementation with a plasmid-encoded *tarIJLM*2 copy (Fig. 2b,c; Extended Data Fig. 1c). The E73 Δ*tarIJLM2* mutant (E73Δ*tarIJLM2*) showed the same growth behavior and microscopic appearance as the parental strain (Extended Data Fig. 4), indicating that the genes have no essential roles in basic cellular processes. Moreover, it had similar phosphate amounts in its cell wall as the parental strain, indicating that the mutant cells probably replaced RboP-WTA by increasing the level of GroP-WTA.

Three large collections of *S. epidermidis* isolates from skin (293 isolates) and nasal cavity (384 isolates) of healthy humans, or from invasive infection (277 isolates) from Tübingen, Hamburg, and Shanghai were screened for the presence of *tarL2* (Table 1). 24 isolates, exclusively derived from infections, were *tarL2*-positive, reaching a 13.9% prevalence in the infection collection from Tübingen. They belonged to the healthcare-associated clones ST2, ST5, ST23, ST87[1,2,16-20], and a new single-locus ST23 variant. In contrast, none of the nasal or skin *S. epidermidis* isolates from any of the collections was *tarL2*-positive indicating that the additional gene cluster may be associated with poor colonization but strong invasion capacities.

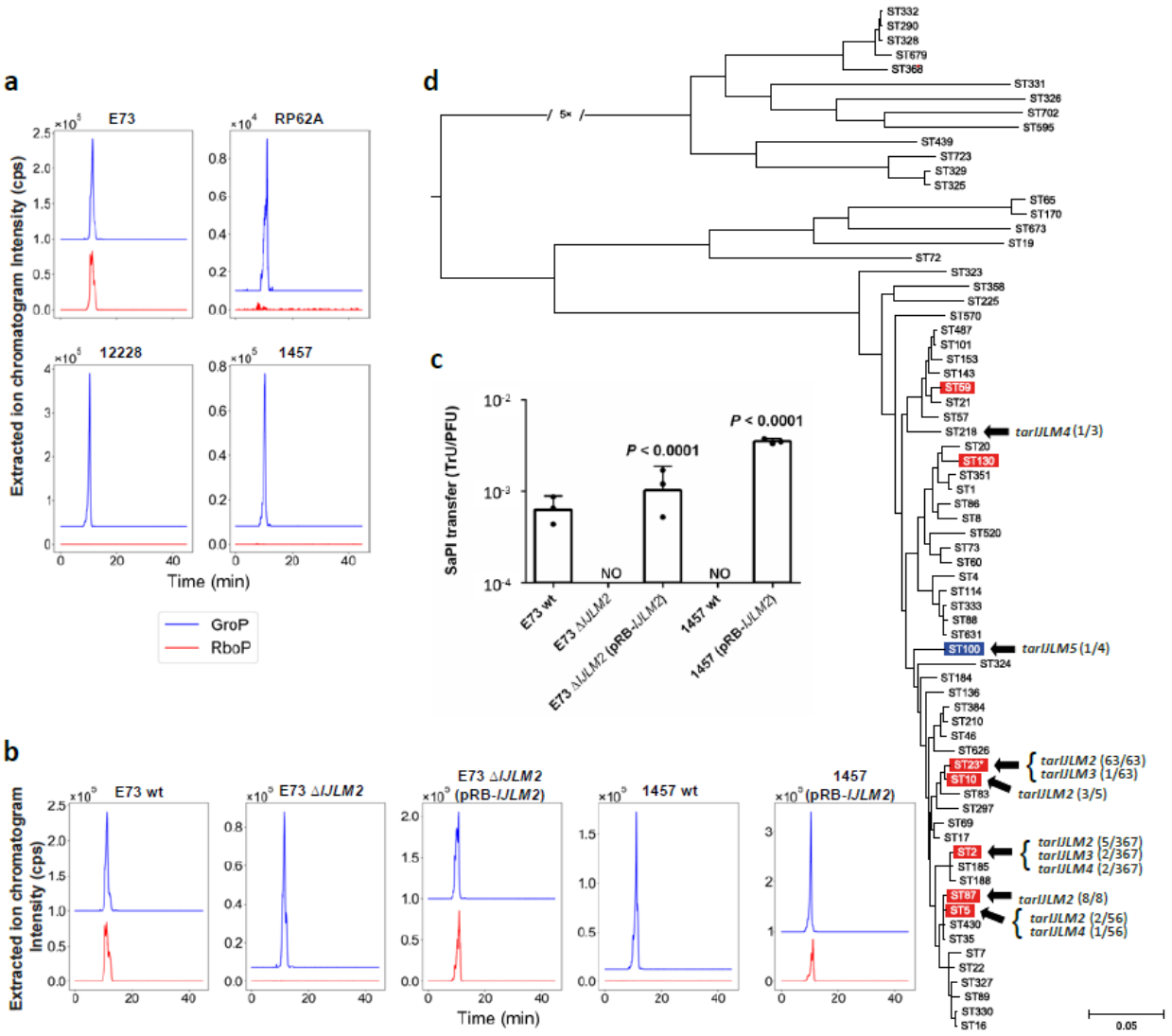


Fig. 2. *tarIJLM* genes are responsible for RboP synthesis in *S. epidermidis* and they are found in healthcare-associated *S. epidermidis* clones. (a) All tested *S. epidermidis* strains contain GroP (blue lines, $[M+H]^+ = 173.022$) but only the *tarIJLM2*-positive E73 has also RboP (red line, $[M+H]^+ = 233.042$) as WTA constituents; counts per second (CPS) measured by liquid chromatography mass spectrometry (LC-MS) are shown. (b) Disruption of *tarIJLM2* abolishes RboP-WTA production in strain E73 while transformation of strain 1457 with the *tarIJLM2*-expressing plasmid pRB-*IJLM2* confers RboP production. (c) Expression of *tarIJLM2* enables *S. epidermidis* to accept SaPIbov1 from transducing *S. aureus* phage Φ11. Values represent means ±

s.d. of three independent experiments. Significant differences vs. E73 wild type (wt) ($P \leq 0.05$) were calculated by one-way ANOVA with Dunnett's post-test (two-sided). (d) *tarIJLM2-5* are almost exclusively found in *S. epidermidis* clones that are predominantly healthcare associated (red) or livestock associated (blue)[21] and have probably been repeatedly imported by HGT. Since the definition of *S. epidermidis* healthcare-association is ambiguous, only clones implicated with the healthcare system in at least three independent studies[1,2,16-20] were labeled red. ST218 has also sporadically been found to cause infections[19]. The maximum-likelihood phylogeny of 71 *S. epidermidis* isolates representing major STs was inferred from an alignment of 87,048 core genome single-nucleotide polymorphisms. The isolates were selected from 497 *S. epidermidis* genomes present in the NCBI Reference Sequence Database (accessed July 3 2018), a global collection of 227 *S. epidermidis* isolates originating from 96 healthcare institutions across 24 countries, and 25 *tarIJLM*-positive *S. epidermidis* isolates collected in this study (Extended Data Table 2). The number of isolates carrying different variants of *tarIJLM* (Extended Data Table 1) and the total number of genomes within each ST (Extended Data Table 2) are indicated. *, includes a new single-locus ST23 variant. The length of the broken branch was reduced 5-fold. The tree was midpoint rooted. The scale bar denotes substitutions per variable sites.

Among 724 publicly available *S. epidermidis* genomes, *tarIJLM2* occurred exclusively in healthcare-associated ST2, ST23, and ST87 isolates but was not detected in any of the 306 isolates from other, predominantly commensal STs (Fig. 2d). Notably, three other *tarIJLM2*-related clusters were found in some human ST2, ST5, ST23, and ST218 genomes, including an ST2 isolate from the infection collection from Tübingen (*tarIJLM3* or *tarIJLM4*), and in a bovine isolate of ST100, which has been associated with lifestock infections[21] (*tarIJLM5*) (Fig. 2d, Extended Data Table 1). The *tarIJLM* genes from the different clusters shared 66-89% nucleotide identities and were more similar to the *S. aureus tarL* genes than the *S. epidermidis tarL1* gene (Extended Data Fig. 5b), suggesting that they may have been acquired from *S. aureus*. The patchy distribution of *tarIJLM* variants in the *S. epidermidis* phylogenetic tree (Fig. 2d, Extended Data Fig. 5a) supports multiple introductions into *S. epidermidis* through several independent HGT events. Consistent with this view, *tarIJLM3* and *tarIJLM4* were located on mobile genetic elements, a plasmid and a SCC element (Fig. 1b), and *tarIJLM*-related genes were also sporadically found in other CoNS species (Extended Data Table 1, Extended Data Fig. 5b, and Supplementary Results).

The absence of *tarIJLM2*-positive *S. epidermidis* from human skin and nose raised the question if RboP-WTA may impair the capacity of *S. epidermidis* to bind to the nasal epithelium, considering that the WTA structure governs nasal colonization by *S. aureus*[22]. Indeed, E73Δ*tarIJLM2* bound 83% better to human airway epithelial A549 cells than the parental strain (Fig. 3a), suggesting that GroP-WTA may be more capable of binding to epithelial receptors than RboP-WTA. Similarly, the GroP-WTA-expressing *S. epidermidis* 1457 bound significantly better to A549 than the same strain expressing *tarIJLM*2. E73Δ*tarIJLM*2 exhibited also a 2.1-fold increased nasal colonization capacity compared to the parental strain in a mouse model (Fig. 3b), which may explain the absence of *tarIJLM*2-positive isolates from the human nose.

Since WTA composition has been found to shape the ability of *S. aureus* to adhere to endothelial cells[23], it was tempting to assume that RboP-WTA is overrepresented among invasive *S. epidermidis* to promote endothelial cell interaction. In contrast to airway epithelial cells, human endothelial cells were significantly less effectively bound by E73Δ*tarIJLM2* compared to the parental and complemented mutant strains (Fig. 3c), suggesting that RboP-WTA may promote the ability of *S. epidermidis* to persist in the bloodstream. Notably, 3.8-fold lower numbers of E73Δ*tarIJLM2* cells were found in the bloodstream of intravenously infected mice compared to the parental strain (Fig. 3d). Moreover, mice infected with the mutant exhibited significantly reduced mortality compared to the wild-type strain (Fig. 3e). *tarIJLM2* augmented virulence *most* likely in a biofilm-independent fashion because E73 and E73Δ*tarIJLM2* lacked biofilm-forming capacities (Extended Data Fig. 4d). Thus, RboP-WTA renders *S. epidermidis* more virulent in invasive infections, probably by increasing its capacity to interact with blood vessel walls.

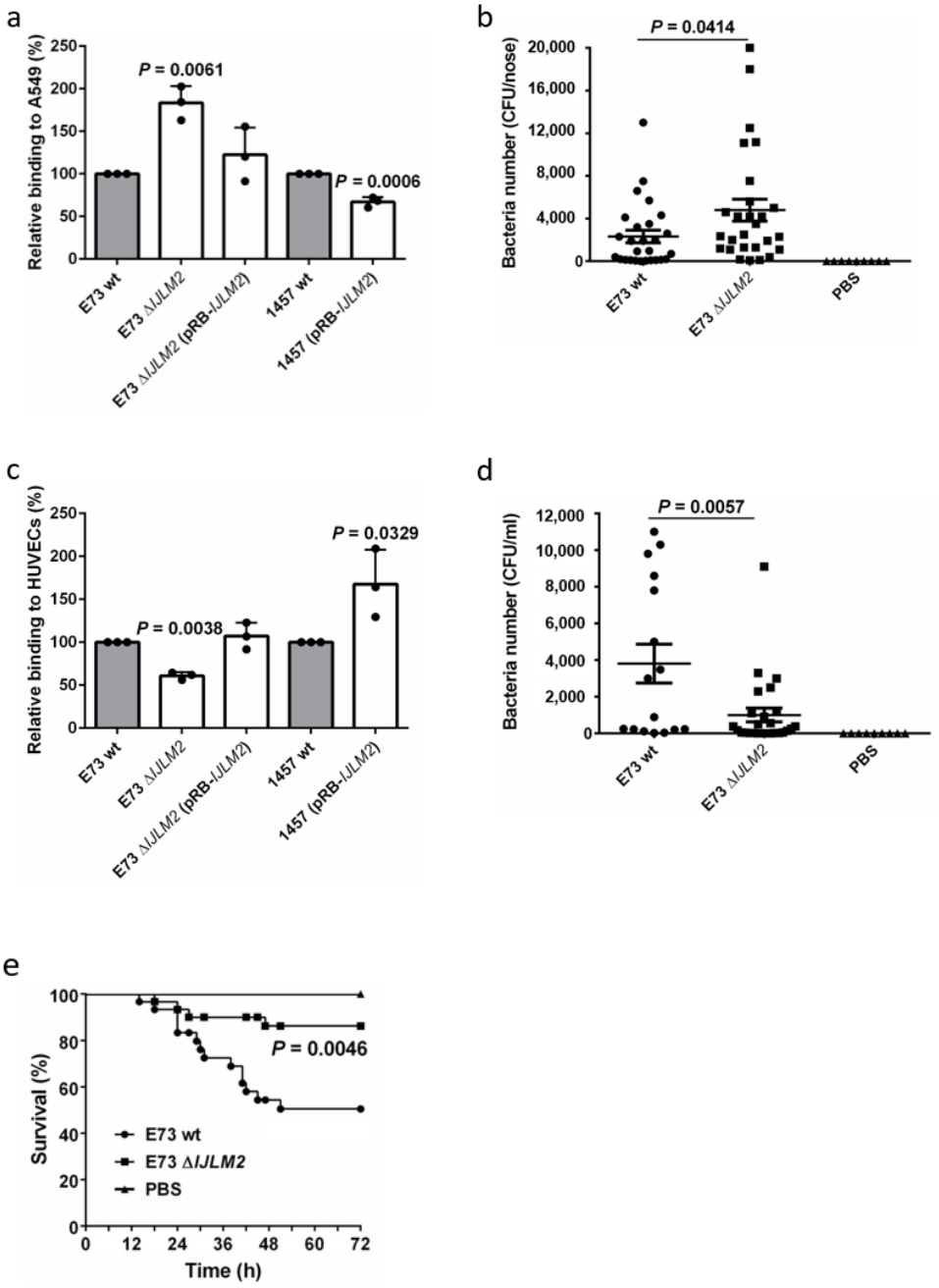


Fig. 3. RboP-WTA impairs colonization but increases infection capacities of *S. epidermidis*. (a) RboP-WTA expression decreases *S. epidermidis* binding to human airway epithelial A549 cells. (b) The E73Δ*tarIJLM2* mutant colonized mouse noses significantly better than the wild type (wt). (c) RboP-WTA expression promotes *S. epidermidis* binding to HUVECs. (d) Blood of

E73Δ*tarIJLM2*-infected mice contains lower CFUs compared to wt-infected mice at 72 h post infection in a bacteremia model. (e) Strain E73 is attenuated for virulence upon deletion of *tarIJLM2* as shown in survival curves from a mouse bacteremia model. Means and s.d. of three independent experiments are shown in (a) and (c). Significant differences ($P \leq 0.05$) were calculated by one-way ANOVA with Dunnett's post-test (two-sided) in (a) and (c) (vs. wt), by unpaired *t*-test (b) and (d), or log-rank (Mantel-Cox) test (e).

WTA is essential for efficient attachment of epithelial cells and nasal colonization by *S. aureus*[24], by binding to scavenger receptors that interact with zwitterionic bacterial polymers such as WTA[22]. However, the diversity and specificity of scavenger receptors on nasal epithelial and potentially also on endothelial cells has remained elusive[25]. Our study indicates that the WTA structure is crucial for *S. epidermidis* interaction with human epithelial and endothelial cells and that RboP- and GroP-WTA differ in their impact on the adhesive properties for different human cell types, thereby shaping the colonization and invasion capacities of this opportunistic pathogen. RboP-WTA is essential for *S. aureus* nasal colonization in addition to several surface-anchored proteins[24,26]. Our study suggests that GroP-WTA may promote nasal colonization even more efficiently than RboP-WTA, which is reflected by the higher prevalence of commensal *S. epidermidis* in the human nose (almost 100%) compared to *S. aureus* (ca. 30% of the population)[27,28].

Future infection control regimes will rely on biomarkers in pathogen genomes allowing the stratification of risks associated with colonization or infection by a specific pathogen clone[29]. The *tarIJLM* cluster represents the first *S. epidermidis* determinant with a strong link to invasiveness and virulence and should therefore be included as an indicator in surveillance programs of this globally important pathogen. It is worth noting that a *tarIJL* cluster is also found in an epidemic lineage of the opportunistic pathogen *Staphylococcus capitis* causing severe infection outbreaks among newborns[30]. RboP-WTA represents a promising antigen for future preventive or therapeutic vaccines that should not only help to fight MRSE infections but may also prevent the phage-dependent HGT of critical resistance, colonization, and virulence factors between *S. epidermidis* and *S. aureus*.

References

1 Lee, J. Y. H. *et al.* Global spread of three multidrug-resistant lineages of Staphylococcus epidermidis. *Nat Microbiol* **3**, 1175-1185 (2018).

2 Li, M., Wang, X., Gao, Q. & Lu, Y. Molecular characterization of Staphylococcus epidermidis strains isolated from a teaching hospital in Shanghai, China. *J Med Microbiol* **58**, 456-461 (2009).

3 Meric, G. *et al.* Disease-associated genotypes of the commensal skin bacterium Staphylococcus epidermidis. *Nat Commun* **9**, 5034 (2018).

4 Heilmann, C., Ziebuhr, W. & Becker, K. Are coagulase-negative staphylococci virulent? *Clin Microbiol Infect*, S1198-743X(18)30739-0. (2018).

5 Weidenmaier, C. & Peschel, A. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nat Rev Microbiol* **6**, 276-287 (2008).

6 Meric, G. *et al.* Ecological Overlap and Horizontal Gene Transfer in Staphylococcus aureus and Staphylococcus epidermidis. *Genome Biol Evol* **7**, 1313-1328 (2015).

7 Moller, A. G., Lindsay, J. A. & Read, T. D. Determinants of Phage Host Range in Staphylococcus Species. *Appl Environ Microbiol* **85**, e00209-19 (2019).

8 Chen, J. *et al.* Genome hypermobility by lateral transduction. *Science* **362**, 207-212 (2018).

9 Lee, A. S. *et al.* Methicillin-resistant Staphylococcus aureus. *Nat Rev Dis Primers* **4**, 18033 (2018).

10 Penades, J. R. & Christie, G. E. The Phage-Inducible Chromosomal Islands: A Family of Highly Evolved Molecular Parasites. *Annu Rev Virol* **2**, 181-201 (2015).

11 Li, M. *et al.* MRSA epidemic linked to a quickly spreading colonization and virulence determinant. *Nat Med* **18**, 816-819 (2012).

12 Winstel, V. *et al.* Wall teichoic acid structure governs horizontal gene transfer between major bacterial pathogens. *Nat Commun* **4**, 2345 (2013).

13 Brown, S., Santa Maria, J. P., Jr. & Walker, S. Wall teichoic acids of gram-positive bacteria. *Annu Rev Microbiol* **67**, 313-336 (2013).

14 Krismer, B. *et al.* Nutrient limitation governs Staphylococcus aureus metabolism and niche adaptation in the human nose. *PLoS Pathog* **10**, e1003862 (2014).

15 Xia, G. *et al.* Glycosylation of wall teichoic acid in Staphylococcus aureus by TarM. *J Biol Chem* **285**, 13405-13415 (2010).

16 Miragaia, M. *et al.* Comparison of molecular typing methods for characterization of Staphylococcus epidermidis: proposal for clone definition. *J Clin Microbiol* **46**, 118-129 (2008).

17 Sivadon, V. *et al.* Partial atlE sequencing of Staphylococcus epidermidis strains from prosthetic joint infections. *J Clin Microbiol* **47**, 2321-2324 (2009).

18 Mertens, A. & Ghebremedhin, B. Genetic determinants and biofilm formation of clinical Staphylococcus epidermidis isolates from blood cultures and indwelling devises. *Eur J Microbiol Immunol (Bp)* **3**, 111-119 (2013).

19 Sharma, P. *et al.* Multilocus Sequence Typing for Interpreting Blood Isolates of Staphylococcus epidermidis. *Interdiscip Perspect Infect Dis* **2014**, 787458 (2014).

20 Post, V. *et al.* Comparative Genomics Study of Staphylococcus epidermidis Isolates from Orthopedic-Device-Related Infections Correlated with Patient Outcome. *J Clin Microbiol* **55**, 3089-3103 (2017).

21 Vasileiou, N. G. C. *et al.* Antimicrobial Agent Susceptibility and Typing of Staphylococcal Isolates from Subclinical Mastitis in Ewes. *Microb Drug Resist*, doi: 10.1089/mdr.2019.0009 (2019).

22 Baur, S. *et al.* A nasal epithelial receptor for Staphylococcus aureus WTA governs adhesion to epithelial cells and modulates nasal colonization. *PLoS Pathog* **10**, e1004089 (2014).

23 Weidenmaier, C. *et al.* Lack of wall teichoic acids in Staphylococcus aureus leads to reduced interactions with endothelial cells and to attenuated virulence in a rabbit model of endocarditis. *J Infect Dis* **191**, 1771-1777 (2005).

24 Weidenmaier, C. *et al.* Role of teichoic acids in Staphylococcus aureus nasal colonization, a major risk factor in nosocomial infections. *Nat Med* **10**, 243-245 (2004).

25 Schade, J. & Weidenmaier, C. Cell wall glycopolymers of Firmicutes and their role as nonprotein adhesins. *FEBS Lett* **590**, 3758-3771 (2016).

26 Weidenmaier, C. *et al.* Differential roles of sortase-anchored surface proteins and wall teichoic acid in Staphylococcus aureus nasal colonization. *Int J Med Microbiol* **298**, 505-513 (2008).

27 Zipperer, A. *et al.* Human commensals producing a novel antibiotic impair pathogen colonization. *Nature* **535**, 511-516 (2016).

28 Iwase, T. *et al.* Staphylococcus epidermidis Esp inhibits Staphylococcus aureus biofilm formation and nasal colonization. *Nature* **465**, 346-349 (2010).

29 Ladner, J. T., Grubaugh, N. D., Pybus, O. G. & Andersen, K. G. Precision epidemiology for infectious disease control. *Nat Med* **25**, 206-211 (2019).

30 Simoes, P. M. *et al.* Single-Molecule Sequencing (PacBio) of the Staphylococcus capitis NRCS-A Clone Reveals the Basis of Multidrug Resistance and Adaptation to the Neonatal Intensive Care Unit Environment. *Front Microbiol* **7**, 1991 (2016).

Methods

Bacterial strains and growth conditions. *S. epidermidis* laboratory strains RP62A, ATCC12228, and 1457, were used as controls for WTA-analytical and phenotypical experiments. *S. aureus* strains RN4220 and PS187 were phage propagation and test strains for phage binding and transduction experiments. *S. aureus* JP1794[31] and PS187-H VW1[12] were used as donor strains for SaPI particle propagation as described below. *E. coli* DC10B was used as a cloning host. RN4220 and PS187Δ*sauUSI*Δ*hsdR*[32] were used as donor strains for plasmid transduction. *S. epidermidis* and *S. aureus* strains were cultivated in tryptic soy broth (TSB) medium or Mueller-Hinton broth (MHB), or otherwise noted and incubated at 37°C on an orbital shaker. *E. coli* strains were cultivated in lysogeny broth (LB). Resistant strains were cultivated in media supplemented with appropriate antibiotics (tetracycline (5 μg ml^{-1}), chloramphenicol (10 μg ml^{-1}), or ampicillin (100 μg ml^{-1})). For experiments performed on solid medium, TSB agar plates containing 5% sheep blood were used unless otherwise noted. A semi-quantitative biofilm assay was performed using 96-well tissue culture plates as previously described with the following modifications[33]. After the cells were fixed in Bouin's fixative for 1 hour, the cells were washed gently three times in PBS and then stained with 0.1% crystal violet solution. The stain was washed off gently three times with PBS and the plates were dried at room temperature. Then 5% acetic acid was added to the wells to dissolve the stain. The absorbance of each well was measured at 570 nm using a CLARIOstar Microplate autoreader (BMG Labtech). *S. epidermidis* isolates were obtained from clinical specimen or from nasal (anterior nasal mucosa of both nostrils) or skin (elbow pit of both arms) swabs from healthy volunteers. Only the infection isolates detected in two independent samples from the same infection sites were taken in this study in order to exclude the contamination *S. epidermidis* isolates from skin in some extent. Bacteria were plated on blood agar plates immediately and directly after swabbing. After overnight incubation, colonies were

analyzed by matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MS) for species determination.

SaPI transfer and phage adsorption assays. Phage propagation was performed according to standard procedures[12]. Briefly, approximately 8.0×10^7 cells of recipient strain grown overnight were mixed with 100 µl of SaPI lysates obtained from *S. aureus* strain JP1794 or PS187-H VW1 bearing the tetracycline resistance marker-labelled SaPIbov1 (~1.0×10^6 PFU ml^{-1}), incubated for 15 min at 37°C, diluted, and plated on TSB agar supplemented with appropriate antibiotics. The adsorption efficiency of Φ11 and Φ187 was determined as described previously with minor modifications[12]. Briefly, adsorption rates were analyzed using a multiplicity of infection (MOI) of 0.1. The adsorption rate was elucidated by determining the number of unbound phages in the supernatant and dividing the number of bound phages by the number of input phages.

Whole-genome sequencing. Whole-genome sequences were generated for the 25 *tarIJLM*-positive *S. epidermidis* isolates collected in this study (Table 1, Extended Data Table 1). Bacterial isolates were incubated overnight on 5% blood agar and bacterial samples were lyzed either by treatment with 10 units/ml lysostaphin (Sigma-Aldrich) and 20 mg/ml lysozyme (Sigma-Aldrich) or mechanically by a bead-beating homogenizer (Precellys). DNA was extracted and quantified using the DNeasy Blood & Tissue Kit (Qiagen) and the Qubit 3.0 Fluorometer (Invitrogen), respectively. Library preparation was conducted in accordance with the Nextera XT DNA Library Prep Guide (Illumina) or NEBNext Ultra library prep Kit for Illumina (New England Biolabs) after shearing the DNA to fragments of 300 bp on a Bioruptor Pico instrument (Diagenode). The libraries were sequenced on a MiSeq platform (Illumina) with 2 × 251 bp using a MiSeq Reagent Kit v2 or on a NextSeq Platform with 2x150 bp using a NextSeq 500/550 v2 kit. Velvet[34] or SPAdes was used to generate *de-novo* assemblies. *S. epidermidis* E73 was also sequenced at the Norwegian Sequencing Centre using the PacBio Sequel platform. DNA was sheared to 10-kb

fragments using g-TUBEs (Covaris). SMRTbell libraries were constructed using standard procedures (Pacific Biosciences). The genome was *de-novo* assembled using the Hierarchical Genome Assembly Process (HGAP4) workflow in SMRT Link version 5.1.0.26411 (Pacific Biosciences). A second *de-novo* assembly was made based on Illumina sequencing and the combination of the sequence contigs resulted in the final genome, which was confirmed by mapping the Illumina reads to the final assembly. The circular chromosome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline.

Sequence analyses. The *tarIJLM2* genes in E73 (accession no. CP035643) were used as queries in BLASTN searches against 497 *S. epidermidis* genomes present in the NCBI Reference Sequence Database (accessed July 3 2018), a global collection of 227 *S. epidermidis* isolates originating from 96 healthcare institutions across 24 countries[1], and the 25 *S. epidermidis* isolates collected in this study (Table 1). MLST typing was performed by comparing contigs with the *S. epidermidis* MLST database[35] using an in-house script. SCC*mec* typing was carried out with the SCC*mec*Finder tool[36]. IS elements were classified according to transposase gene similarity by using BLAST analysis with the ISfinder database[37].

Phylogenetic analyses. Mapping of sequence reads and SNP calling were carried out using the Northern Arizona SNP Pipeline (NASP)[38]. Sequence reads were mapped against the *S. epidermidis* ST2 reference isolate BPH0662 (GenBank accession no. NC_017673) using the Burrows-Wheeler Alignment tool[39]. SNP calling was achieved using the GATK Unified Genotyper[40,41] with the following parameters: ≥10× mapping coverage; ≥90% unambiguously base calls; insertions and deletions were ignored. SNPs contained in repeats, as determined by NUCmer[42,43], were excluded. Phages in BPH0662 were identified with PHASTER[44] and SNPs residing within these were manually removed from the alignment, whereafter Gubbins[45] was used to remove recombination tracts. The MUSCLE algorithm[46] was used to construct multiple

sequence alignments for each of the *tarIJLM* genes and calculate pairwise nucleotide identities. A subset of *tarIJLM* genes representing the diversity observed within *S. epidermidis* was used as queries in BLASTN searches against non-*S. epidermidis* genomes present in the NCBI Reference Sequence Database (accessed October 29 2018). Phylogenetic reconstruction was carried out using the maximum-likelihood program PhyML with a GTR model of nucleotide substitution[47,48]. Support for the nodes was assessed using aBayes[49].

Specific PCR screening for different types of *tarL* genes. Isolates were screened for different variants of *tarL* genes, by PCR amplification using the primers listed in Extended Data Table 3. The thermocycler conditions for the amplification were: 98°C 2 min, 98°C 30 s, 55°C 30 s, 72°C 1 min, 72°C 10 min. The PCR products were confirmed by sequencing (Eurofins, Germany).

Cell wall and WTA isolation. Cell walls and WTA were isolated as described previously[24]. Briefly, bacteria were grown overnight in TSB supplemented with 0.25% glucose. Bacterial cells were disintegrated with a FastPrep-24 instrument (MP Biomedicals). Lysates were incubated overnight with DNase I (40 units ml^{-1}, Roche) and RNase A (80 units ml^{-1}, Sigma) at 37°C. Cell wall and WTA samples were dialysed against pyrogen-free water (Ambion). WTA amounts were quantified by phosphate assay as described previously[24]. To quantify the WTA amount per cell, cell walls from each test strain were isolated as described above. A volume of 300 µl cell wall suspension was mixed with 300 µl 1 M NaOH and incubated at 60°C with constant shaking of 600 rpm in a Thermomixer (Eppendorf). After 2 h, the phosphate content in supernatants of this mixture was measured as described above. The same amount of 300 µl cell wall was dried in a Speedvac concentrator and weighted to determine the phosphate amount per cell wall dry mass.

WTA chromatographic and MS analysis. The composition of pure WTA samples was determined by methanolysis of the samples with 0.5 M HCl, MeOH at 85°C for 45 min followed by twice acetylation using acetic anhydride and pyridine (1:1, v/v) at 85°C for 10 min and

detection by gas-liquid chromatography (GLC) and GLC-MS. The systems used were an Agilent Technologies 7890A instrument equipped with a HP-5MS capillary column (30 m x 0.25 mm, film thickness 0.25 µm) and applying a temperature gradient of 150°C (kept for 3 min) to 250°C at 3°C/min for GLC or a Agilent Technologies 7890A instrument equipped with a dimethylpolysiloxane column (Agilent, HP Ultra 1, 12 m x 0.2 mm, film thickness 0.33 µm) and 5975C series MSD detector with electron impact ionization (EI) mode under autotuned condition at 70 eV applying a temperature gradient of 70°C (kept for 1.5 min) to 110°C at 60°C/min and then to 320°C at 5°C/min. For the WTA sample of E73 wild type, the absolute configuration was determined by GLC by comparison with authentic standards of the acetylated (S)-2-butanol glycoside derivative after butanolysis (2 M HCl, (S)-2-butanol at 85°C for 2 h) and acetylation as described before. The same system was used for detection applying a temperature gradient of 120°C (kept for 3 min) to 320°C at 3°C/min.

Identification of WTA polymer type was performed using an Ultimate 3000RS HPLC system (Dionex) coupled to a micrOTOFII electrospray-ionization time-of-flight mass spectrometer (Bruker). For HPLC, a Gemini C18 column (150 x 4.6 mm, 110 Å, 5 µM, Phenomenex) was used at 37°C with a flow rate of 0.2 ml/min. A 5-min equilibration step with 100% buffer A (0.1% formic acid, 0.05% ammonium formate) was applied, followed by a linear gradient of 0 to 40% buffer B (acetonitrile) for 30 min. A final washing step with 40% buffer B for 5 min and a re-equilibration step (100% buffer A) for 5 min completed the method. Samples were ionized via electrospray ionization (ESI) in positive ion mode. Exact masses in positive ion mode for GroP $[M+H]^+$ = 173.022 m/z and RboP $[M+H]^+$ = 233.042 m/z were presented as extracted ion chromatograms with Data Analysis (Bruker). Base peak chromatograms were used for sample normalization.

Nuclear magnetic resonance (NMR) spectroscopy experiments were carried out in D_2O at 300

K. All one-dimensional (1H and 13C) and two-dimensional homo- (COSY, TOCSY, and ROESY) and heteronuclear (HSQC-DEPT, HMBC) experiments were recorded with a Bruker AvanceII 700 MHz spectrometer (operating frequencies of 700.43 MHz for 1H NMR, and 176.13 MHz for 13C NMR) using standard Bruker software. COSY, TOCSY, and ROESY were recorded using data sets (t1 by t2) of 4,096 by 512 points, and 4 scans (E73 wild-type, 1457 wild-type, and 1457 (pRB-*IJLM2*)), 2 scans (E73 Δ*tarIJLM2*) or 1 scan (E73 Δ*tarIJLM2* (pRB-*IJLM2*)) were acquired for each t1 value in the case of COSY, 8 scans (E73 wild type, E73 Δ*tarIJLM2*, E73 Δ*tarIJLM2* (pRB-*IJLM2*), and 1457 (pRB-*IJLM2*)) or 16 scans (1457 wild type) in the case of TOCSY and 8 scans in the case of ROESY. The TOCSY experiment was carried out in the phase-sensitive mode with mixing times of 120 ms. The 1H,13C correlations were measured in the 1H-detected mode via HSQC-DEPT with proton decoupling in the 13C domain acquired using data sets of either 4,096 by 512 points (16 scans for each t1 value in the case of E73 Δ*tarIJLM2* and E73 Δ*tarIJLM2* (pRB-*IJLM2*); 48 scans for each t1 in the case of 1457 wild type and 1457 (pRB-IJLM2) or 4,096 by 1,024 points (32 scans for each t1 in the case of E73 wild type). The HMBC spectra were acquired using data sets of 4,096 by 512 points and either 32 scans (E73 Δ*tarIJLM2* and 1457 wild type), 36 scans (E73Δ*IJLM2*(pRB-*IJLM2*)), 104 scans (E73 wild type) or 128 (1457(pRB-*IJLM2*)) for each t1 value. Chemical shifts are reported relative to external acetone (1H 2.225; 13C 31.50).

Molecular genetic methods. For the construction of the Δ*tarIJLM2* mutant in *S. epidermidis* E73, the pBASE6-erm/lox1 shuttle vector was used according to standard procedures[50]. For mutant complementation experiments, plasmid pRB474 was used. The primers for knockout and complementation plasmid construction are listed in Extended Data Table 3. Plasmid transduction to *S. epidermidis* strains was performed using □11 with *S. aureus* RN4220 as donor strain or Φ187 as transduction phage with *S. aureus* PS187 as donor strain according to the method by

Winstel *et al*[32]. Briefly, RN4220 strains bearing knockout or complementation plasmid were infected with Φ11 and the lysates were used to infect recipient strains. To this end, overnight-grown bacteria were resuspended in phage buffer (100 mM $MgSO_4$, 100 mM $CaCl_2$, 1 M Tris–HCl, pH 7.8, 0.59% NaCl, 0.1% gelatine), mixed with 100 µl lysate (~1×10^8 PFU/ml), and incubated at 37°C for 10 min. The mixture was then plated on TSB agar containing 12.5 mg/ml tetracycline.

RNA isolation and qRT-PCR. Bacterial cultures grown overnight in TSB were diluted in fresh TSB, PBS with 50% pooled heat-inactivated human serum from healthy donors, or synthetic nasal medium (SNM3, containing 0.2 mM bipyridine)[14] and grown at 37°C and 180 rpm. Bacteria were harvested after 6 h (exponential growth phase) and immersed in RNAprotect bacteria reagent (Qiagen). Bacterial cells were then washed and resuspended in RLT buffer, followed by a mechanical lysis with a bead-beating homogenizer (Precellys). RNA was extracted from the lysate with the RNeasy Mini Kit (Qiagen). RNA was transcribed into cDNA with iScript reverse transcriptase (bio-rad), and qRT-PCR was performed on a LightCycler480 instrument (Roche). Transcription levels of *tarL1* and *tarL2* were normalized against the expression of the constitutively transcribed *gyrB, rho,* and *tpiA* housekeeping genes (Extended data table 3). PCR set-up, cycling conditions and assays for housekeeping genes were as described previously[51]. Each experiment was performed in biological and technical duplicates. qPCR results were analyzed in LightCycler 480 SW 1.5.1. and GenEx software (multiD).

FITC labeling of *S. epidermidis* cells. Overnight *S. epidermidis* cultures were diluted in fresh MHB and grown at 37°C to log growth phase (OD 600 nm = 1.0). Bacteria were then harvested, washed three times with PBS, and labeled with fluorescein isothiocyanate (FITC; 0.1 mg/ml) for 1 h at 37°C. The bacteria were washed three times with PBS and resuspended in RPMI medium (Sigma). The concentration of FITC-labeled bacteria was determined using a Neubauer chamber.

FITC-labeled bacteria was stored at -80°C.

Cell culture. A549 human bronchial epithelial cells (American Type Culture Collection (ATCC CCL-185)) were purchased from ATCC and grown in DMEM-F12 medium (Gibco-BRL) supplemented with 100 µg/ml streptomycin, 100 U/ml penicillin, 10% heat-inactivated fetal bovine serum, and 2 mM glutamine. Human umbilical vein endothelial cells (HUVECs) cryopreserved from pooled donors were purchased from PromoCell and grown in Endothelial Cell Growth Medium (PromoCell). All cells used here were grown at 37°C under 5% CO_2. A549 and HUVECs were used until passage nine and seven, respectively.

Interaction of *S. epidermidis* with epithelial or endothelial cells. The capacity of *S. epidermidis* to adhere to epithelial cells was analyzed as described previously[24]. Briefly, approximately 1×10^6 A549 or HUVEC cells were seeded to ibiTreat slides (Ibidi) to form confluent cell monolayers. Cells were grown at 37°C under 5% CO_2 in appropriate cell culture media. The cells in slides were washed twice with DMEM (A549) or DMEM-F12 (HUVECs) medium and inoculated with FITC-labeled bacteria for 1 h at 37°C under 5% CO_2. Multiplicities of infection (MOI) with FITC-labeled bacteria of 100 were used. To mimic the liquid flow in nose and blood stream, peristaltic pumps (Amersham) were used during infection (30 min) with FITC-labeled bacteria at a flow rate of 8.4 ml/h at 0.3 dynes/cm^2. The cells' nuclei and cytoskeletons were stained with 4',6-diamidino-2-phenylindole and phalloidin-tetramethylrhodamine B isothiocyanate, respectively, and fixed using 3.5% paraformaldehyde (PFA)-PBS at room temperature. PFA was removed, and wells were coated with 1 ml PBS. No morphological changes of bacterial or human cells were observed after this procedure. Bacteria adhering to epithelial or endothelial cells were counted using a confocal microscope.

Nasal colonization model. Female four to six-weeks old BALB/c mice were used for the nasal colonization model (27 mice per bacterial strain). Mice were given water containing kanamycin

(25 μg/ml) for three days before infection. Bacterial cells were grown to mid-exponential growth phase (OD 600 nm = 1.0), washed, and resuspended in PBS. The inoculum, which contained 10^{10} CFU in 10 μl PBS or PBS alone, was pipetted slowly onto the nares of anesthetized mice without touching the skin with the pipette tip. Three days after inoculation, mice were euthanized, nasal regions were wiped externally with 70% ethanol, and the nasal tissues were homogenized in 0.5 ml of TSB. The total number of *S. epidermidis* CFU per nose was assessed by plating 100 μl diluted nasal suspensions on TSA with 5% sheep blood and enumerating colonies after overnight growth at 37°C.

Bacteremia model. Female four to six-weeks old BALB/c mice were used for the bacteremia model (30 mice per bacterial strain). Bacterial cells were grown in TSB to mid-exponential growth phase (OD 600 nm = 1.0), washed, and resuspended in sterile PBS. Anesthetized mice received 10^{11} CFUs of E73 wild type or mutant in 0.1 ml PBS, or PBS alone by retro-orbital injection via the right eye. After inoculation, general animal health condition and disease advancement were monitored regularly for up to 72 h. Evaluation of animal morbidity was based on the following criteria: hunched posture, activity levels, ruffled fur, and labored breathing, etc. All surviving mice were euthanized at 72 h post infection. The blood samples were plated on TSA with 5% sheep blood to measure CFU of *S. epidermidis* in blood.

Statistical analysis. Statistical analysis was performed using the Prism 6.0 package (GraphPad Software). P values of ≤ 0.05 were considered significant.

Ethics statement. Bacterial samples from human volunteers were collected upon written informed consent and approval by the institutional review boards of the Universities of Tübingen (577/2015A) and Shanghai Jiaotong University (protocol 2017001). Isolates from Hamburg were obtained in accordance with §12 of the Hamburg hospital law (HmbKHG). All mouse work was performed in China, which was approved by the Ethics Committee of Renji Hospital, School of

Medicine, Shanghai Jiaotong University.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability. All major data generated or analyzed in this study are included in the article or its supplementary information files. The genome sequences of *S. epidermidis* isolates were deposited at GenBank; the accession numbers can be found in Extended Data Table 1. Source data for experiments with animals (Fig 3b, d, and e) are provided. All other data relating to this study are available from the corresponding author on reasonable request.

Extended data

Extended Data Table 1 | List of *tarIJLM*-positive *S. epidermidis* and other CoNS isolates

Strain	Collection	Year	Country	Host	Source	MLST	*mecA*	*tarIJLM*	Accession no.
S. epidermidis									
IRL01	Lee *et al*.[1]	2016	Republic of Ireland	Human	Bacteraemia	2	+	2	SAMN09103952
US06	Lee *et al*.	2010	United States	Human	Prosthetic joint infection	5	+	4	SAMN09103969
AUS27	Lee *et al*.	2017	Australia	Human	Colonising isolate	23	+	2	SAMN09093632
AUS28	Lee *et al*.	2017	Australia	Human	Colonising isolate	23	-	2	SAMN09093633
AUT01	Lee *et al*.	2016	Austria	Human	Unknown	23	+	2	SAMN09103856
BEL02	Lee *et al*.	2015	Belgium	Human	Bacteraemia	23	+	2	SAMN09103858
BEL03	Lee *et al*.	2015	Belgium	Human	Bacteraemia	23	+	2	SAMN09103859
BEL05	Lee *et al*.	2016	Belgium	Human	Bacteraemia	23	+	2	SAMN09103861
BEL17	Lee *et al*.	2017	Belgium	Human	Bacteraemia	23	+	2	SAMN09103873
BPH0666	Lee *et al*.	2012	Australia	Human	Bacteraemia	23	+	2	ERS1019850
BPH0695	Lee *et al*.	2007	Australia	Human	Bacteraemia	23	+	2	ERS1019876
BPH0703	Lee *et al*.	2007	Australia	Human	Bacteraemia	23	+	2	ERS1019883
BPH0713	Lee *et al*.	2008	Australia	Human	Bacteraemia	23	+	2	ERS1019893
BPH0715	Lee *et al*.	2008	Australia	Human	Bacteraemia	23	-	2	ERS1019895
DEN02	Lee *et al*.	2001	Denmark	Human	Unknown	23	+	2	SAMN09103876
DEN09	Lee *et al*.	2001	Denmark	Human	Unknown	23	+	2	SAMN09103883
DEN11	Lee *et al*.	2001	Denmark	Human	Unknown	23	+	2	SAMN09103885
DEN12	Lee *et al*.	2012	Denmark	Human	Colonising isolate	23	+	2	SAMN09103886
DEN13	Lee *et al*.	2012	Denmark	Human	Colonising isolate	23	+	2	SAMN09103887
DEN14	Lee *et al*.	2013	Denmark	Human	Colonising isolate	23	+	2	SAMN09103888
DEN16	Lee *et al*.	2014	Denmark	Human	Colonising isolate	23	+	2	SAMN09103890
DEN18	Lee *et al*.	2015	Denmark	Human	Colonising isolate	23	+	2	SAMN09103892
DEN19	Lee *et al*.	2015	Denmark	Human	Colonising isolate	23	+	2	SAMN09103893
DEN20	Lee *et al*.	2016	Denmark	Human	Central venous catheter	23	+	2	SAMN09103894
DEN21	Lee *et al*.	2016	Denmark	Human	Central venous catheter	23	+	2	SAMN09103895
DEN22	Lee *et al*.	2016	Denmark	Human	Colonising isolate	23	+	2	SAMN09103896
DEN24	Lee *et al*.	2016	Denmark	Human	Colonising isolate	23	+	2	SAMN09103898

DEN25	Lee *et al*.	2016	Denmark	Human	Central venous catheter	23	+	2	SAMN09103899
DEN26	Lee *et al*.	2016	Denmark	Human	Central venous catheter	23	+	2	SAMN09103900
DEN27	Lee *et al*.	2016	Denmark	Human	Central venous catheter	23	+	2	SAMN09103901
DEN29	Lee *et al*.	2017	Denmark	Human	Central venous catheter	23	+	2	SAMN09103903
DEN30	Lee *et al*.	2017	Denmark	Human	Central venous catheter	23	+	2	SAMN09103904
DEN31	Lee *et al*.	2017	Denmark	Human	Colonising isolate	23	+	2	SAMN09103905
DEN32	Lee *et al*.	2017	Denmark	Human	Colonising isolate	23	+	2	SAMN09103906
DEN33	Lee *et al*.	2017	Denmark	Human	Colonising isolate	23	+	2	SAMN09103907
DEN34	Lee *et al*.	2017	Denmark	Human	Colonising isolate	23	+	2	SAMN09103908
DEN35	Lee *et al*.	2017	Denmark	Human	Colonising isolate	23	+	2	SAMN09103909
DEN36	Lee *et al*.	2017	Denmark	Human	Skin and Soft Tissue Infection	23	+	2	SAMN09103910
DEN37	Lee *et al*.	2017	Denmark	Human	Colonising isolate	23	-	2	SAMN09103911
FRA01	Lee *et al*.	2005	France	Human	Bacteraemia	23	+	2	SAMN09103913
FRA07	Lee *et al*.	2010	France	Human	Septic arthritis	23	+	2	SAMN09103919
FRA14	Lee *et al*.	2015	France	Human	Bone and joint infection	23	+	2	SAMN09103926
FRA16	Lee *et al*.	2004	France	Human	Bacteraemia	23	+	2	SAMN09103928
FRA17	Lee *et al*.	2004	France	Human	Bacteraemia	23	+	2	SAMN09103929
GER08	Lee *et al*.	2015	Germany	Human	Bacteraemia	23	+	2	SAMN09103938
GER15	Lee *et al*.	2016	Germany	Human	Bacteraemia	23	+	2	SAMN09103945
GER19	Lee *et al*.	2017	Germany	Human	Osteomyelitis	23	+	2	SAMN09103949
US02	Lee *et al*.	2000	USA	Human	Prosthetic joint infection	23	+	2	SAMN09103966
US03	Lee *et al*.	2004	USA	Human	Prosthetic joint infection	23	+	2	SAMN09103967
US04	Lee *et al*.	2007	USA	Human	Prosthetic joint infection	23	+	2	SAMN09103976
US07	Lee *et al*.	2012	USA	Human	Prosthetic joint infection	23	+	2	SAMN09103970
US09	Lee *et al*.	2014	USA	Human	Prosthetic joint infection	23	+	2	SAMN09103972
DEN23	Lee *et al*.	2016	Denmark	Human	Colonising isolate	87	+	2	SAMN09103897

UK04	Lee *et al*.	2016	UK	Human	Bacteraemia	87	+	2	SAMN09103956
DAR1907	NCBI	2007	USA	Human	Blood	2	+	3	CP013943, CP016969, CP016970, CP016971
ENVH150	NCBI	2004	Germany	-	-	2	+	3	NZ_LYVW00000000
VCU013	NCBI	-	-	-	-	2	+	4	NZ_JHTZ00000000
SAM-3	NCBI	2016	Lebanon	Human	Sputum	23	+	2	NZ_PYFN00000000
VCU126	NCBI	-	-	-	-	23	+	2	NZ_AHLG00000000
VCU117	NCBI	-	-	-	-	23	+	2 + 3	NZ_AHLA00000000
C3	NCBI	2013	Denmark	Human	Eye	87	+	2	NZ_OAOC00000000
SNUC_2569	NCBI	2007	Canada	Cattle	Herd 302	100	-	5	NZ_PYYU00000000
NIHLM003	NCBI	2008	USA	Human	Umbilicus	218	-	4	NZ_AKHB00000000
RJ4	This study	2016	China	Human	Catheter	2	+	2	SAMN11658252
RJ5	This study	2016	China	Human	Catheter	2	+	2	SAMN11658253
RJ6	This study	2016	China	Human	Blood	2	+	2	SAMN11658254
RJ7	This study	2016	China	Human	Catheter	2	+	2	SAMN11658255
RJ8	This study	2016	China	Human	Cerebrospinal fluid	2	+	4	SAMN11658256
E51	This study	2014	Germany	Human	Ulcer	5	+	2	SAMN11658242
E56	This study	2014	Germany	Human	Wound	5	+	2	SAMN11658243
E22	This study	2014	Germany	Human	Wound	10	+	2	SAMN11658238
PIA099	This study	2015	Germany	Human	Joint	10	+	2	SAMN11658248
E6	This study	2014	Germany	Human	Central venous catheter	23	+	2	SAMN11658236
E13	This study	2014	Germany	Human	Central venous catheter	23	+	2	SAMN11658237
E50	This study	2014	Germany	Human	Wound	23	+	2	SAMN11658241
E73	This study	2014	Germany	Human	Wound	23	+	2	CP035643
HD43	This study	2016	Germany	Human	Joint	23	+	2	CP040867
PIA017	This study	2015	Germany	Human	Joint	23	+	2	SAMN11658246
PIA062	This study	2015	Germany	Human	Joint	23	+	2	SAMN11658247
E1	This study	2014	Germany	Human	Wound	87	+	2	SAMN11658235
E45	This study	2014	Germany	Human	Wound	87	+	2	SAMN11658240
E94	This study	2014	Germany	Human	Wound	87	+	2	SAMN11658244
HD33	This study	2016	Germany	Human	Joint	87	+	2	CP040864
HD66	This study	2017	Germany	Human	Joint	87	+	2	CP040868
RJ1	This study	2014	China	Human	Abscess	SLV of ST23	+	2	SAMN11658249
RJ2	This study	2014	China	Human	Abdominal dropsy	SLV of ST23	+	2	SAMN11658250
RJ3	This study	2014	China	Human	Catheter	SLV of ST23	+	2	SAMN11658251
S. hominis									
SNUC_3404	NCBI	2007	Canada	Cattle	Herd 323	NA	-	5 (-*tarM*)	NZ_PZHY00000000

SNUC_5746	NCBI	2008	Canada	Cattle	Herd 324	NA	-	5 (-*tarM*)	NZ_PZHW00000000
SNUC 5748	NCBI	2008	Canada	Cattle	Herd 324	NA	-	5 (-*tarM*)	NZ_QXVQ00000000
S. warneri									
SNUC_194	NCBI	2007	Canada	Cattle	Herd 102	NA	-	3	NZ_PZFE00000000
Staphylococcus sp.									
HMSC061F01	NCBI	-	-	Human	Vaginal/ rectal	NA	+	5	NZ_LTNM01000000
HMSC069E07	NCBI	-	-	Human	Vaginal	NA	+	4 (-*tarM*)	NZ_LTPB01000000

MLST, multilocus sequence type; SLV, single-locus variant; NA, not applicable.

Extended Data Table 2 | List of *S. epidermidis* genomes representing 71 multilocus sequence types

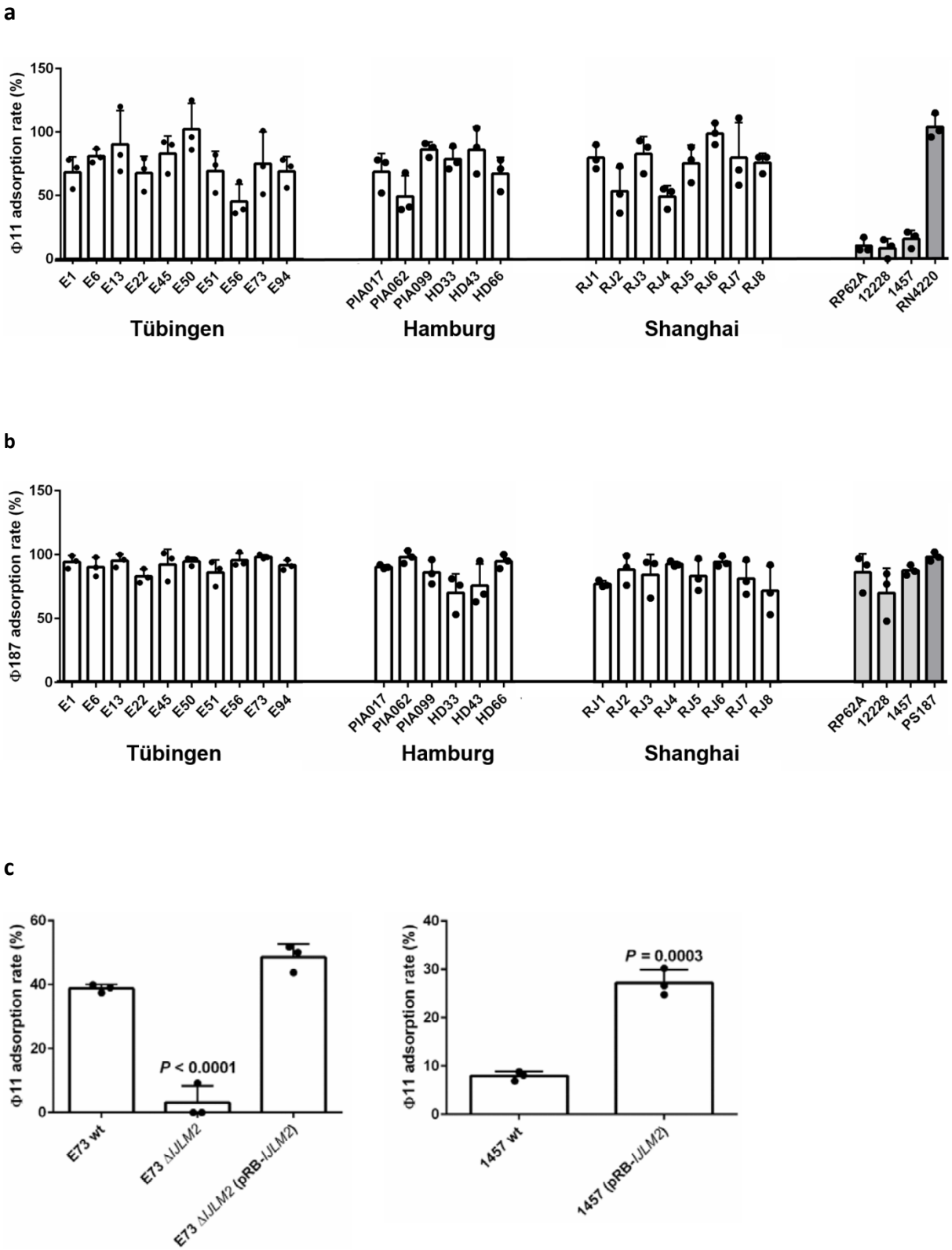
MLST	Total no. of isolates	Representative strain	Collection	Accesion no.
1	2	VCU081	NCBI	NZ_AHLU01000000
2	367	BPH0662	Lee *et al*.[1]	NZ_LT571449
4	1	VCU036	NCBI	NZ_JHUA01000000
5	56	BPH0723	Lee *et al*.	ERS1019903
7	8	BPH0690	Lee *et al*.	ERS1019871
8	3	ATCC12228	Lee *et al*.	ERS1019847
10	5	RP62A	Lee *et al*.	ERS1019921
16	5	US11	Lee *et al*.	SAMN09103974
17	5	1296_SEPI	NCBI	NZ_JVVA00000000
19	1	VCU071	NCBI	NZ_AGUB00000000
20	4	FRA09	Lee *et al*.	SAMN09103921
21	1	US10	Lee *et al*.	SAMN09103973
22	6	FRA02	Lee *et al*.	SAMN09103914
23	63	BPH0666	Lee *et al*.	ERS1019850
35	4	SCH-22	NCBI	NZ_PEKK02000000
46	1	BPH0707	Lee *et al*.	ERS1019887
57	2	SNUC 1209	NCBI	NZ_PYZC00000000
59	25	BPH0667	Lee *et al*.	ERS1019851
60	1	644_SEPI	NCBI	NZ_JUYV00000000
65	1	NLAE-zl-G239	NCBI	NZ_FOPD00000000
69	5	91_SEPI	NCBI	NZ_JUOJ00000000
72	2	NIHLM087	NCBI	NZ_AKGK00000000
73	9	NIHLM039	NCBI	NZ_AKGS01000000
83	7	SCH-12	NCBI	NZ_PEJB00000000
86	4	BPH0700	Lee *et al*.	ERS1019880
87	8	DEN23	Lee *et al*.	SAMN09103897

88	4	SCH-20	NCBI	NZ_PEJJ00000000
89	3	BPH0697	Lee *et al*.	ERS1019878
100	4	SNUC_2569	NCBI	NZ_PYYU00000000
101	2	SCH-11	NCBI	NZ_PEJA00000000
114	1	8400	NCBI	NZ_JMIF00000000
130	9	SCH-17	NCBI	NZ_PEJG00000000
136	1	VCU127	NCBI	NZ_AHLH00000000
143	1	FRA08	Lee *et al*.	SAMN09103920
153	2	Scl25	NCBI	NZ_ATDC02000000
170	2	NIHLM040	NCBI	NZ_AKGR00000000
184	4	PM221	Lee *et al*.	HG813242
185	1	NIH051668	NCBI	NZ_AKHK00000000
188	2	BPH0693	Lee *et al*.	ERS1019874
210	4	BPH0699	Lee *et al*.	ERS1019879
218	3	NIHLM003	NCBI	NZ_AKHB00000000
225	2	BPH0677	Lee *et al*.	ERS1019862
290	3	SCH-15	NCBI	NZ_PEJE00000000
297	2	SNUC 2400	NCBI	NZ_PYYY00000000
323	1	NIHLM001	NCBI	NZ_AKHC00000000
324	1	NIHLM008	NCBI	NZ_AKHA00000000
325	1	NIHLM015	NCBI	NZ_AKGZ00000000
326	1	NIHLM031	NCBI	NZ_AKGX00000000
327	3	1340_SEPI	NCBI	NZ_JVSY00000000
328	1	NIHLM023	NCBI	NZ_AKGU00000000
329	3	NIHLM037	NCBI	NZ_AKGT00000000
330	1	NIHLM049	NCBI	NZ_AKGQ00000000
331	1	NIHLM057	NCBI	NZ_AKGO00000000
332	1	NIHLM061	NCBI	NZ_AKGN00000000
333	1	NIHLM067	NCBI	NZ_AKGM00000000
351	1	MIT 14-1777-C6	NCBI	NZ_NXMC00000000
358	1	SCH-23	NCBI	NZ_PEKJ00000000
368	1	BPH0704	Lee *et al*.	ERS1019884
384	1	VCU125	NCBI	NZ_AHLF00000000
430	1	NIH04003	NCBI	NZ_AKHJ00000000
439	1	NGS-ED-1109	NCBI	NZ_JZUL00000000
487	4	BPH0684	Lee *et al*.	ERS1019866
520	1	971_SEPI	NCBI	NZ_JULN00000000
570	2	SNUC 2450	NCBI	NZ_PYYX00000000
595	1	UC7032	NCBI	NZ_ARWU00000000
626	1	PEI-B-P-06	NCBI	NZ_FUVD00000000
631	1	643_SEPI	NCBI	NZ_JUYW00000000
673	1	SNUT	NCBI	NZ_LQRB00000000
679	1	ATCC12228?	NCBI	NZ_CP022247
702	10	CIM40	NCBI	NZ_ATCW02000000
723	2	785_SEPI	NCBI	NZ_JUTA01000000
ND	58	-	-	-

The list includes 749 *S. epidermidis* genomes from the NCBI Reference Sequence Database (accessed July 3 2018), a global collection of hospital isolates described by Lee *et al*.[1], and 25 *tarIJLM*-positive isolates collected in this study (Table 1).
MLST, multilocus sequence type; ND, not determinable by the existing MLST scheme.

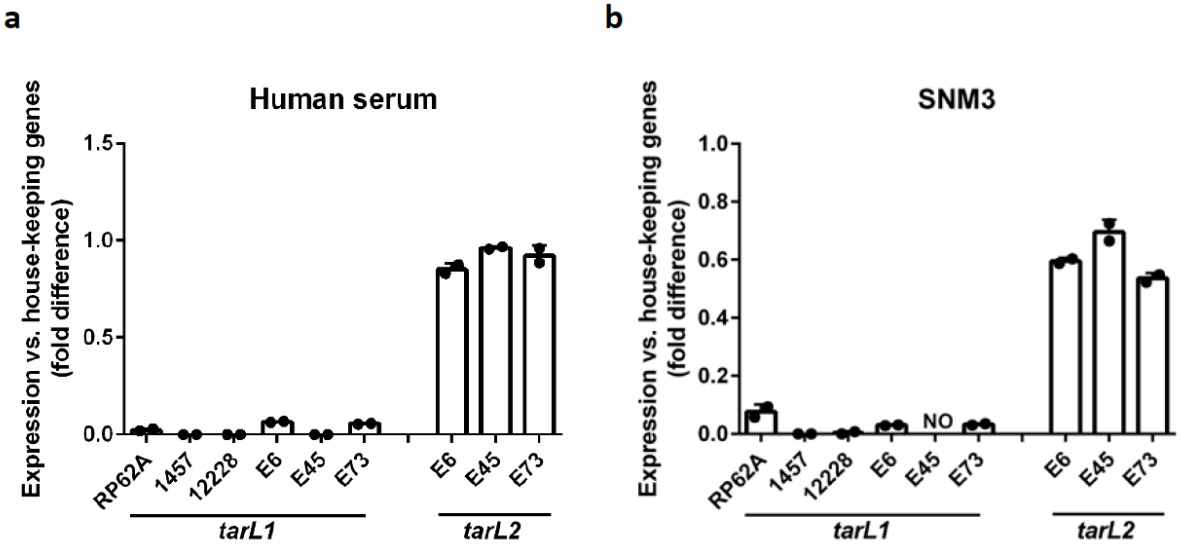
Extended Data Table 3. Oligonucleotides used in this study.

Name	Sequence
For *tarL* screening PCR	
tarL1-up	GACTTTAGAGGAGTTAAGAGATTACAACTCAC
tarL1-dn	GAAGTTTATATATCATCGGATGATAATCATCAACC
tarL2-up	CTTAACGTTTCGGTTAGGATACCAACTC
tarL2-dn	CAACTGTGTTATTACTTTCACAATAACTTGC
tarL3-up	GTGGTGGTTAAAGGAAATAATTTTAAATGTAG
tarL3-dn	CCATTTATAATTGGCAATAATTGTTCCATTC
tarL4-up	CTTGTAGATTTAACGTTGCAATATTAGATAATG
tarL4-dn	GGTTTAAATAAACTGTGGATTTTATATTTATCATC
For *tarIJLM2* knock out	
IJLM-1up BASE	CATGAGATCTCCTCTATTTGTTTCAGTGTAGAACGATCATGG
IJLM-1dn BASE	CTTCTTCCTCATTATCCTTCTCTAAGTATTGC
IJLM-2up BASE	GAAGGATAATGAGGAAGAAGTAAAGTATCCATTTTTCTATTATTCAGTTTCTATG
IJLM-2dn BASE	GTCAGTCGACACCATAAATCCTGTTCCCATACTATC
For *tarIJLM2* complementation	
IJLM2_F	GTCGGATCCAAAGGAGGTTATATAATGAAATATGCTGGAATACTTGCGGG
IJLM2_R	CCGATGAATTCTTAGTTAAACAATTGTTTCCATTTCACCATCA
For qRT-PCR	
tarL_1_f	TGCTAATCGTGTTAATATTCAAGGATC
tarL_1_r	GTTGTAATCTCTTAACTCCTCTAAAGTC
tarL_1_probe	FAM-ACCGACAAGAAGATGAATACGTCTGCCA-BHQ1
tarL_2_f	CGAAAGCATTTGTGTCATCAGCA
tarL_2_r	CCCCTAAAAGTTGGCGCAAACA
tarL_2_probe	FAM- GAGCAGAATGTAATTCCAACAGGAGTACCT-BHQ1

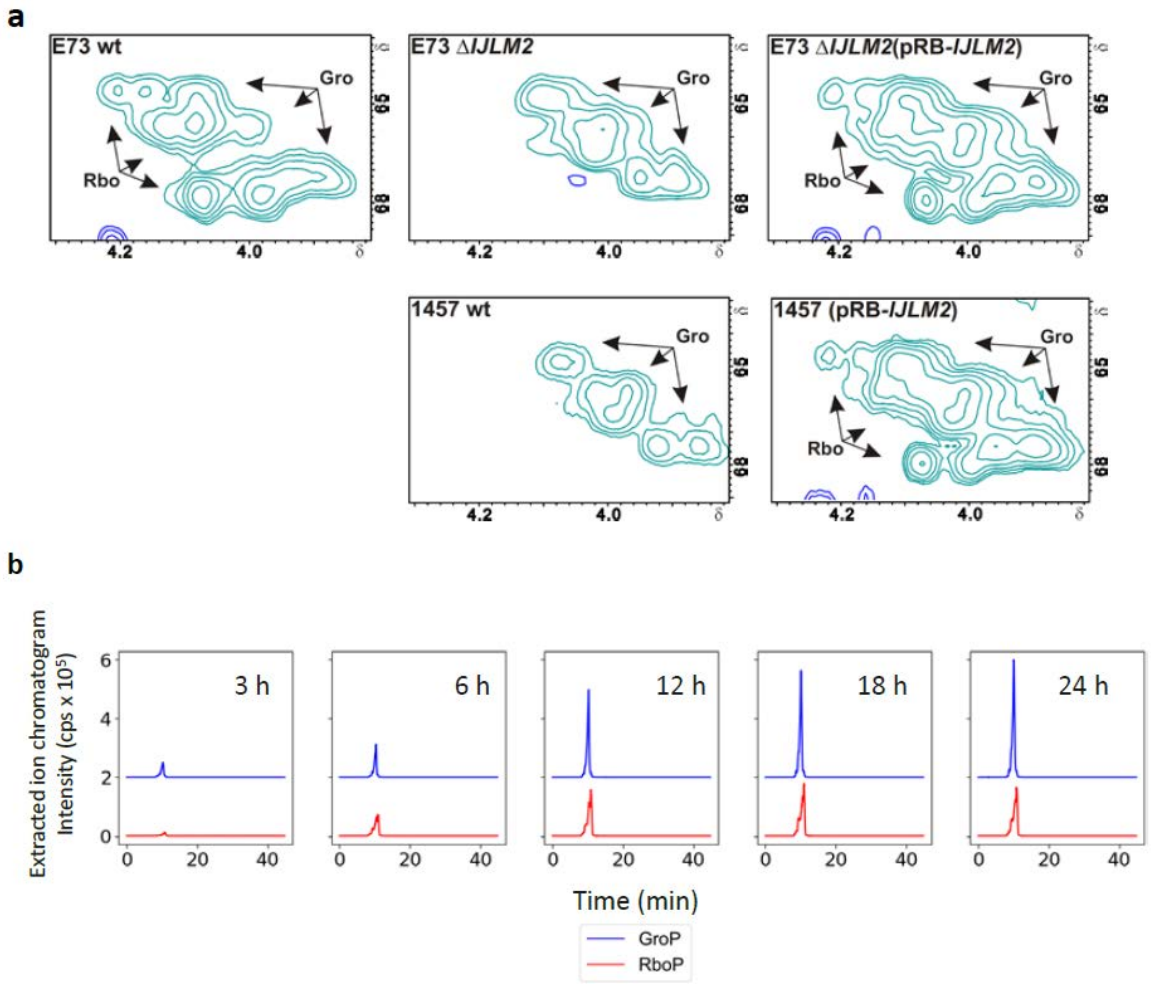


Extended Data Fig. 1. Phage adsorption assay confirms presence of the RboP-WTA phage

receptor on *tarIJLM2*-positive *S. epidermidis* isolates. The binding rate of RboP-WTA-specific phage Φ11 (a, c) and of GroP-WTA specific Φ187 (b) to the indicated strains is shown. RP62A, ATCC12228, and 1457 are *S. epidermidis* laboratory strains (light gray), only expressing GroP-WTA. *S. aureus* strains RN4220 or PS187 (dark gray) express RboP-WTA or GroP-WTA, respectively. Means and s.d. of three independent experiments are shown. Student‘s t-test was used to analyze significant differences vs. wt in (c).

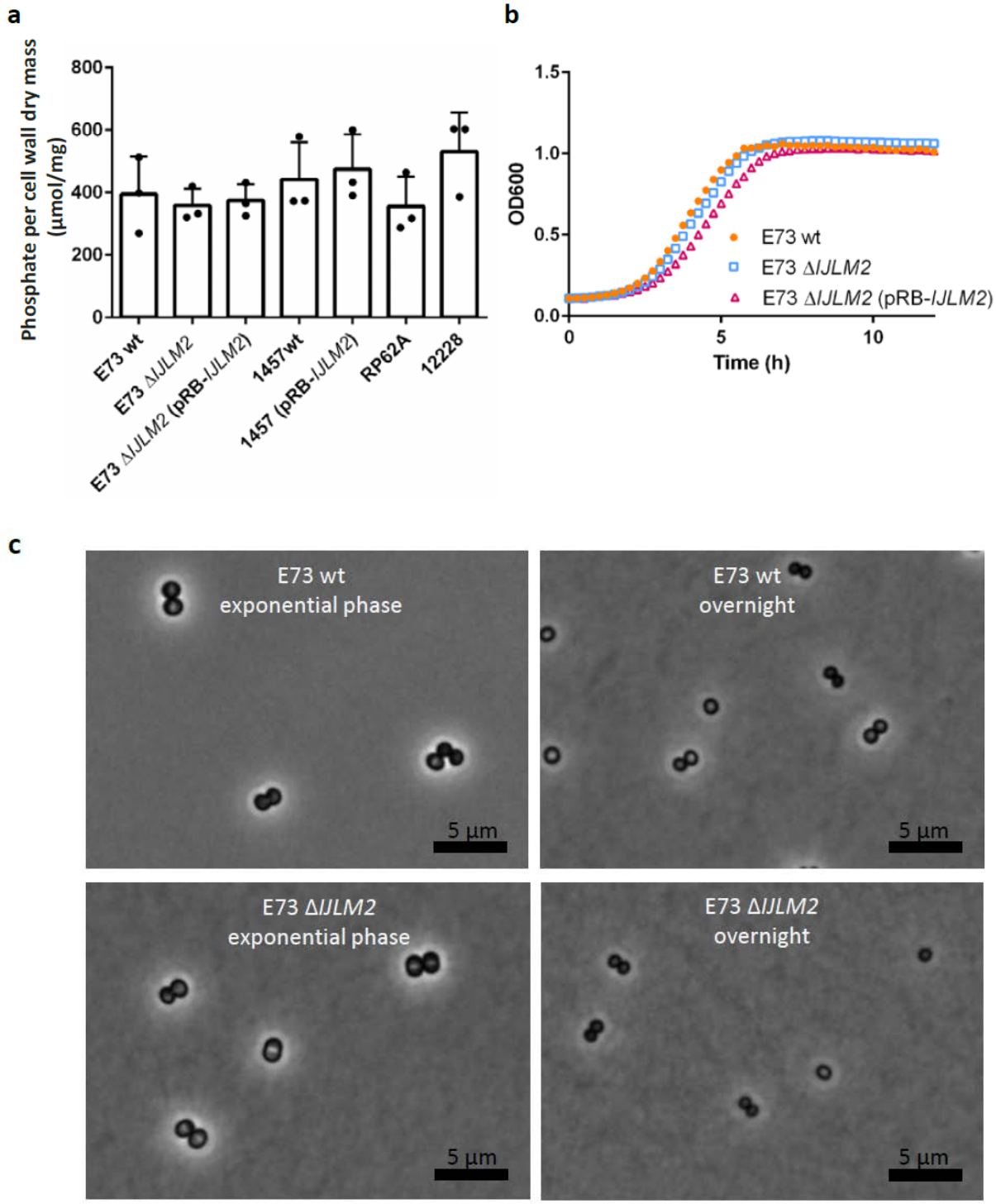


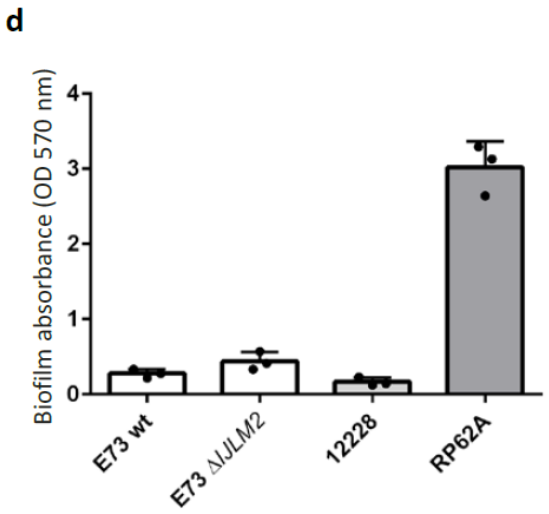
Extended Data Figure 2. While *tarL2* is efficiently transcribed in the *tarIJLM2*-positive isolates E6, E45, and E73, *tarIJL1* is not or only very weakly expressed in all tested *S. epidermidis* strains during growth in human serum (a) or synthetic nasal medium 3 (SNM3) (b). Values represent means ± s.d. of two independent experiments. They were normalized for strongly and constitutively expressed housekeeping genes *gyrB*, *rho*, and *tpiA*. NO, below detection limit.



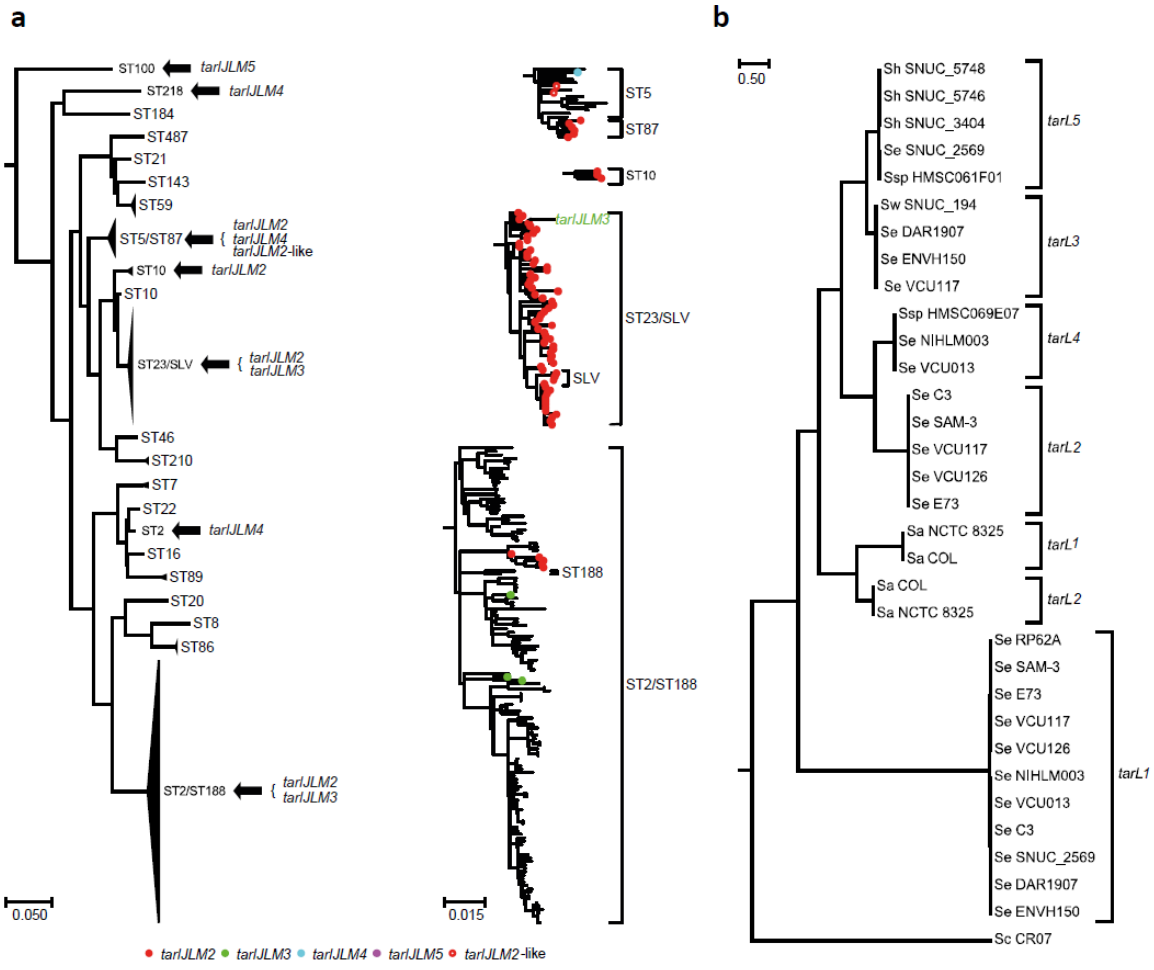
Extended Data Fig. 3. Nuclear magnetic resonance (NMR) and MS spectroscopy confirm that WTA from *tarIJLM2*-positive *S. epidermidis* strains contains both, glycerol (Gro) and ribitol (Rbo), while WTA of *tarIJLM2*-lacking strains contains only Gro. (a) Part of the HSQC-DEPT spectra of the methylene region of Rbo/Gro in the poly-RboP/GroP of the indicated WTA samples are shown. Signals corresponding to different substitutions of the polymers are shifted depending on the substitution. The presence of poly-RboP is clearly seen in the case of the WTA samples E73 wild type (wt), E73Δ*IJLM2* (pRB-*IJLM2*) and 1457 (pRB-*IJLM2*) with signals at d 4.13;4.09/65.7, 4.07;3.98/67.8 and 4.20;4.16/64.7 (H/C), which are absent in the case of the other WTA samples. (b) *S. epidermidis* E73 produces both, GroP (blue lines, $[M+H]^+$ =

173.022 m/z) and RboP (red line, $[M+H]^+$ = 233.042 m/z) in different growth phase as shown by LC-MS.





Extended Data Fig. 4. RboP-WTA expression does not increase the overall WTA amounts and does not affect *S. epidermidis* growth behavior, microscopic appearance, or biofilm formation. (a) Total WTA phosphate amount per cell wall dry weight of the indicated *S. epidermidis* strains. Means and s.d. of three independent experiments are shown, none of the minor differences is significant. (b) Growth curves of *S. epidermidis* E73 with or without *tarIJLM2* in TSB. (c) Microscopy of E73 with or without *tarIJLM2*. (d) Neither E73 wild-type nor E73Δ*IJLM2* forms biofilms. The laboratory *S. epidermidis* strains ATCC12228 (biofilm negative) and RP62A (biofilm positive) were used as control strains.



Extended Data Fig. 5. Detailed phylogenetic distribution of *tarIJLM2-5* among selected *S. epidermidis* clones and evolutionary relation of *tarL* from different gene clusters. (a) The maximum-likelihood phylogeny of 261 genomic sequences, comprising 25 *tarIJLM*-positive *S. epidermidis* isolates collected in this study (Table 1), nine *tarIJLM*-positive *S. epidermidis* genomes from in the NCBI Reference Sequence Database (accessed July 3 2018), and a global collection of 227 *S. epidermidis* isolates originating from 96 healthcare institutions across 24 countries[1] was inferred from an alignment of 39,298 single-nucleotide polymorphisms. Four *tarIJLM*-negative isolates (BPH0677, BPH0704, BPH0737, and SEI) were manually removed from the phylogeny due to their extreme divergence, and clades containing isolates with the same

or closely related STs were collapsed to reduce complexity. The STs are indicated. Expanded subtrees illustrate the phylogenetic relationships of isolates belonging to ST5/ST87, ST10, ST23/SLV, and ST2/ST188. The presence of *tarIJLM2-5* is indicated in colours. A single ST23 isolate, VCU117, carried both, *tarIJLM2* and *tarIJLM3*. (b) The *tarL2* found in *S. epidermidis* is more closely related to the two *tarL* genes in *S. aureus* than to *S. epidermidis tarL1*. The *S. epidermidis tarL3, tarL4, and tarL5* genes, but not *tarL2*, were also found in other CoNS. The maximum-likelihood phylogeny was inferred from *tarL* genes present in *S. epidermidis* (Se) RP62A (*tarIJL*), E73 (*tarIJL* + *tarIJLM*), C3 (*tarIJL* + *tarIJLM*), DAR1907 (*tarIJL* + *tarIJLM*), ENVH150 (*tarIJL* + *tarIJLM*), NIHLM003 (*tarIJL* + *tarIJLM*), SAM-3 (*tarIJL* + *tarIJLM*), SNUC_2569 (*tarIJL* + *tarIJLM*), VCU013 (*tarIJL* + *tarIJLM*), VCU117 (*tarIJL* + *tarIJLM* + *tarIJLM*), VCU126 (*tarIJL* + *tarIJLM*), *S. warneri* (Sw) SNUC_194 (*tarIJLM*), *S. hominis* (Sh) SNUC_3404 (*tarIJL*), SNUC_5746 (*tarIJL*), and SNUC_5748 SNUC (*tarIJL*), uncharacterised *Staphylococcus* species (Ssp) HMSC061F01 (*tarIJL*) and HMSC069E07 (*tarIJLM*), *S. aureus* (Sa) COL (*tarIJL1* + *tarIJL2* [accession no. CP000046]) and NCTC 8325 (*tarIJL1* + *tarIJL2* [NC_007795]), and *S. capitis* (Sc) CR07 (*tarIJL* [NZ_CZWH00000000]). The trees were midpoint rooted and the scale bars denote substitutions per variable sites.

Supplementary Information

Supplementary Results.

The *tarIJLM2* gene cluster was found to be integrated into the same chromosomal region between two conserved genes in all 81 *tarIJLM2*-positive *S. epidermidis* isolates (Fig. 1b). The genes flanking *tarIJLM2* in these isolates were also present in the chromosome of most (161/182) of the *tarIJLM2*-negative *S. epidermidis* isolates, but here they flanked another gene with no homology to *tarIJLM* (Fig. 1b). The absence of mobile genetic elements (e.g., insertion sequences) between *tarIJLM2* and the flanking genes indicates that the missing gene might have been replaced by *tarIJLM2* in a process called homology-facilitated illegitimate recombination[51]. The *tarIJLM3* gene cluster was located on a 29-kb unannotated plasmid in the ST2 isolate DAR1907, and on 23- and 28-kb annotated contigs of unknown genomic origin in the other isolates. These contigs contained the *repA* gene encoding plasmid replication protein and shared 100% nucleotide identities with the plasmid sequence from DAR1907, indicating that they are likely of plasmid origin (Fig. 1b). Of note, the *tarIJLM3* gene cluster was flanked by two pseudogenes homologous to insertion sequence (IS) elements belonging to the IS*6* family (Fig. 1b). The *tarIJLM4* gene cluster was located 1,079 bps downstream of the *orfX* gene in all three isolates (Fig. 1b). Analysis of the contig containing the *tarIJLM4* gene cluster in the ST218 isolate NIHLM003 showed that it was carried on a 38-kb staphylococcal cassette chromosome (SCC) element, which was integrated into a unique site in the 3'-end of *orfX* (referred to as integration site sequence, ISS), contained the *ccrA2B2* genes encoding site-specific recombinases, and was flanked by direct repeat sequences containing the ISS. The contigs containing *orfX* and the *tarIJLM4* gene cluster from isolate VCU013 (ST2) and US06 (ST5) shared 100% nucleotide identities with the first 12 and 8 kbs of the SCC element from NIHLM003, respectively. Apart from the SCC element, the ST2 and ST5 isolates each contained a *mec* and *ccr* gene complex

characteristic of a type III (3A) and type IV (2B) SCC*mec* element, respectively, but they were present on different contigs and it was therefore not possible to determine their location in relation to the SCC element. We were also unable to determine the genomic location of the *tarIJLM5* gene cluster in the bovine ST100 isolate SNUC_2569, due to its presence on two short contigs with no flanking sequences.

The different *tarIJLM* gene variants were used as queries in BLASTN searches against non-*S. epidermidis* genomes present in the NCBI Reference Sequence Database (accessed 3 July 2018). The contigs containing the *tarIJLM3* gene cluster shared 98% identity with a 7-kb region on a 33-kb contig from a bovine *S. warneri* isolate (Extended Data Table 1), including the *tarIJLM3* gene cluster and flanking IS*6*-like pseudogenes. The presence of the *repA* gene on the *S. warneri* contig indicated a plasmid origin. The remaining part of the *S. warneri* contig showed little homology to the *tarIJLM3*-carrying plasmid-like contigs from *S. epidermidis,* supporting mobilisation of the *tarIJLM3* cluster between plasmids from different *Staphylococcus* species. Analysis of the contigs containing the *tarIJLM4* gene cluster showed that the left extremity of the SCC element shared 98% identity with a 5-kb region at the right extremity of a contig from an isolate belonging to a yet uncharacterised *Staphylococcus* species (Extended Data Table 1). The contig contained the left ISS and the *tarIJL4* gene cluster. Of note, *tarM4* was not present in any of the contigs from this isolate. The *ccrA2B2* genes were located on a separate contig. The *tarIJLM5* gene cluster present in the bovine ST100 isolate SNUC_2569 was identified on a 14-kb SCC element in an isolate belonging to another uncharacterised *Staphylococcus* species (Extended Data Table 1). The SCC element contained the *ccrA4B4* genes and was preceded by a 23-kb SCC element, which was integrated into the 3'-end of *orfX* and contained the *ccrA1B1* genes. Besides the *tarIJLM5* gene cluster, the two SCC elements showed little homology to other contigs from SNUC_2569. The *tarIJL5* gene cluster, but not *tarM5*, were also present in SCC

elements in three bovine *S. hominis* isolates (Extended Data Table 1). In contrast to the other *tarIJLM* gene clusters, we did not find evidence for the presence of *tarIJLM2* homologues outside *S. epidermidis*. Nonetheless, these data support the hypothesis that the *tarIJLM* gene clusters were derived from other *Staphylococcus* species through exchange of genetic material. Furthermore, the identification of the *tarIJLM3* and *tarIJLM5* gene clusters in bovine *S. warneri* and *S. epidermidis* isolates, respectively, suggests the existence of an animal reservoir.

References:

31 Tormo, M. A. *et al.* Staphylococcus aureus pathogenicity island DNA is packaged in particles composed of phage proteins. *J Bacteriol* **190**, 2434-2440 (2008).

32 Winstel, V., Kuhner, P., Rohde, H. & Peschel, A. Genetic engineering of untransformable coagulase-negative staphylococcal pathogens. *Nat Protoc* **11**, 949-959 (2016).

33 Wang, L. *et al.* SarZ is a key regulator of biofilm formation and virulence in Staphylococcus epidermidis. *J Infect Dis* **197**, 1254-1262 (2008).

34 Zerbino, D. R. & Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18**, 821-829 (2008).

35 Thomas, J. C. *et al.* Improved multilocus sequence typing scheme for Staphylococcus epidermidis. *J Clin Microbiol* **45**, 616-619 (2007).

36 Kaya, H. *et al.* SCCmecFinder, a Web-Based Tool for Typing of Staphylococcal Cassette Chromosome mec in Staphylococcus aureus Using Whole-Genome Sequence Data. *mSphere* **3**, e00612-17 (2018).

37 Siguier, P., Perochon, J., Lestrade, L., Mahillon, J. & Chandler, M. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* **34**, D32-36 (2006).

38 Sahl, J. W. *et al.* NASP: an accurate, rapid method for the identification of SNPs in WGS datasets that supports flexible input and output formats. *Microb Genom* **2**, e000074 (2016).

39 Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595 (2010).

40 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303 (2010).

41 DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491-498 (2011).

42 Delcher, A. L., Phillippy, A., Carlton, J. & Salzberg, S. L. Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res* **30**, 2478-2483 (2002).

43 Kurtz, S. *et al.* Versatile and open software for comparing large genomes. *Genome Biol* **5**, R12 (2004).

44 Arndt, D. *et al.* PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* **44**, W16-21 (2016).

45 Croucher, N. J. *et al.* Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* **43**, e15 (2015).

46 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797 (2004).

47 Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**, 696-704 (2003).

48 Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-321 (2010).

49 Anisimova, M., Gil, M., Dufayard, J. F., Dessimoz, C. & Gascuel, O. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst Biol* **60**, 685-699 (2011).

50 Geiger, T. *et al.* The stringent response of Staphylococcus aureus and its impact on survival after phagocytosis through the induction of intracellular PSMs expression. *PLoS Pathog* **8**, e1003016 (2012).

51 Weiser, J. *et al.* Sub-inhibitory tigecycline concentrations induce extracellular matrix binding protein Embp dependent Staphylococcus epidermidis biofilm formation and immune evasion. *Int J Med Microbiol* **306**, 471-478 (2016).

Chapter 3

-

A novel *Staphylococcus epidermidis* phage ΦTÜB: the genetic characteristics and its function as a tool for high efficient plasmid transduction

Xin Du[1,2], Tibor Botka[3], Ivana Maslanova[3], Lena Mühlenbruch[1,2], Pavol Bardy[3], Volker Winstel[4], Jesper Larsen[5], Ralf Rosenstein[1,2], Andreas Peschel[1,2].

Affiliations: 1. Infection Biology, Interfaculty Institute of Microbiology and Infection Medicine, University of Tübingen, Germany. 2. German Center for Infection Research (DZIF), partner site Tübingen, Germany. 3. Department of Experimental Biology, Faculty of Science, Masaryk University, Brno, Czech Republic. 4. Institute for Medical Microbiology and Hospital Epidemiology, Hannover Medical School, Hannover, Germany. 5. Statens Serum Institut, Denmark.

To be submitted. Target Journal: *Applied and Environmental Microbiology*.

Abstract

Staphylococcus epidermidis is one of the most common pathogens in various types of nosocomial infections in hospitals. Currently, *S. epidermidis* is drawing increasing attention from researchers for its high resistance to antibiotics, its ability to form biofilms, and the difficulty of distinguishing commensal strains from pathogenic ones. However, it remains difficult to introduce plasmids into *S. epidermidis* strains. To solve this problem, new methods and tools must be discovered. We isolated a new phage from a clinical *S. epidermidis* strain in Tübingen. Here, we report the characteristics of this new phage ΦTÜB and its strong ability to transduce plasmids into *S. epidermidis*. ΦTÜB is a *Siphoviridae* phage with a very long tail, as observed by electron microscopy. ΦTÜB had a narrow host range of only *S. epidermidis*. It could infect more than 40% of the *S. epidermidis* strains in a phage infection assay with both methicillin-sensitive and methicillin-resistant clinical *S. epidermidis* strains. Genome sequencing showed that the ΦTÜB genome consists of 44,592 bp of dsDNA with a GC content of 34.5%. Although ΦTÜB showed similarity with the *Staphylococcus aureus* phage Φ11 in host recognition genes, it could only adsorb to *S. epidermidis* but not to *S. aureus*. ΦTÜB is easily propagated in *S. epidermidis* strain 1457. ΦTÜB showed stability in TSB medium at pH values between 5.0 and 8.0 and temperatures below 50°C. ΦTÜB can transduce plasmid DNA efficiently even to strains refractory to electroporation. Therefore, ΦTÜB might become a valuable research tool for transduction for *S. epidermidis* strains, which are often difficult to transform.

Introduction

Staphylococcus epidermidis is a common commensal bacterium not only on the skin of healthy humans but also on the mucosa of the mouth and nose[1-3]. In recent years, *S. epidermidis* has gained much attention due to its frequency in causing hospital diseases and its importance in promoting intra- and inter-species evolution[4,5]. *S. epidermidis* forms biofilms on medical devices and is a serious clinical threat as one of the major causes of nosocomial infections[6,7]. It has also been reported that *S. epidermidis* causes mastitis in lactating women[8]. It has also been recognized as one of the main aetiological agents of ovine and bovine mastitis threatening animal health[9]. Although *S. epidermidis* itself contains few virulence genes in its genome, evidence has shown that *S. epidermidis* acts as a gene reservoir and a key factor in the transmission of genes for *Staphylococcus aureus*[10]. In addition, *S. epidermidis* has high resistance rates towards many antibiotics, which, together with its biofilm, raises the expectation of searching for new drugs or other treatment methods[11,12].

An essential step for the studies of *S. epidermidis* is the introduction of plasmids into *S. epidermidis* strains. Several methods for plasmid introduction in s*taphylococci* are not efficient for *S. epidermidis*. The first important method is electroporation. However, for unknown biological reasons (e.g., CRISPR, restriction-modification systems), electroporation in *S. epidermidis* is much more difficult than that in *S. aureus*[13,14]. Researchers have tried to improve the efficiency by modifying the electronic voltage, the recovery medium, and preparation methods of competent cells. Most of these have been of little help in increasing the transformation efficiency. The efficiencies of electroporation in *S. epidermidis* are still magnitudes lower than those in *S. aureus*. For *staphylococci*, protoplast transformation is also used[15]. However, it does not work for *S. epidermidis* due to the natural resistance of this bacteria to lysostaphin digestion[16]. The third method, phage transduction, was reported to be efficient by Winstel *et al*[17]. The *S. aureus* phage Φ187 can efficiently transduce plasmids into *S. epidermidis* via glycerolphosphate wall teichoic acid. However, for unknown reasons, not all *S. epidermidis* strains could be transduced using Φ187[17].

Therefore, more phages are needed for plasmid introduction. While there are several studies of *S. aureus* phage transduction for plasmid[18-20], the study of *S. epidermidis* phage transduction is lacking.

However, while numerous studies have focused on *S. aureus* phages, only a few *S. epidermidis* phages have been reported and studied until now. *S. epidermidis* phages were first reported in 1979 for phage typing of s*taphylococci*[21]. Most *S. epidermidis* phages reported today belong to *Siphoviridae*, while only a few are *Myoviridae* or *Podoviridae* phages. The genome sequences of sixteen phages and prophages from *S. epidermidis* are in the NCBI database, while few studies on the isolation and characterization of *S. epidermidis* phages have been reported. The genomes of phage PH15 and phage CNPH82 were the first two completely sequenced phage genomes. They have high sequence homology, belong to the *Siphoviridae* family and produce stable lysogens[22]. Due to a defective lysogeny module, vB_SepiS_phiIPLA5 is strictly lytic[23,24]. The *S. epidermidis* phage vB_SepS_SEP9 is unable to lysogenize because it encodes a nonfunctional integrase and has no recognizable lysogeny module[25]. SEP1 is an *S. epidermidis* myovirus that is highly specific to *S. epidermidis* strains and has a high lytic spectrum, making it a good therapeutic candidate[26]. All *S. epidermidis* genomes are organized with five functional modules: DNA metabolism, DNA packaging, capsid morphogenesis, tail morphogenesis, lysis, and lysogeny.

In this paper, we introduce a novel *S. epidermidis* phage, ΦTÜB, from an *S. epidermidis* strain isolated from an infected tooth of a clinical patient hospitalized in the Hospital of Tübingen University. ΦTÜB has a strong ability to transduce plasmids into a large number of clinical *S. epidermidis* strains. The novel ΦTÜB is a promising genetic tool for *S. epidermidis* study.

Method

Ethics Statement. Bacterial samples from patients were collected upon written informed consent and approval by the institutional review boards of Tübingen

University (013/2014B02 – 015/2014BO2).

Bacterial strains and growth conditions. The bacterial strains collected and used in this study are listed in Table 1. There were an additional 114 *S. epidermidis* strains isolated from the hospital of Tübingen University, of which 77 strains were excluded because of chloramphenicol resistance. In total, 37 clinical *S. epidermidis* strains were used to test the transduction efficiency of ΦTÜB. *S. epidermidis* and *S. aureus* strains were cultivated in tryptic soy broth (TSB) medium and incubated at 37°C with shaking in an orbital shaker at 160 revolutions per minute (rpm). *E. coli* strains were cultivated in LB (Luria Bertani broth) medium. Resistant strains were cultivated in media supplemented with the appropriate antibiotics (chloramphenicol (10 μg ml^{-1}) or ampicillin (100 μg ml^{-1})). Growth characteristics were monitored at 37°C. For experiments performed on solid medium, TSB agar plates containing 5% sheep blood were used unless otherwise noted.

Electron and Cryo-electron microscopy. Purified phage was diluted to a concentration of A280nm=0.8. For negative staining, 4 µl of sample was applied onto glow-discharged carbon-coated 400 mesh copper grids for 2 min. The grids were then washed twice on a drop of deionized water and stained with 2% (w/v) uranyl acetate. For cryo-EM, 3.9 µl of sample was applied onto glow-discharged 400 mesh Quantifoil grids, blotted and plunge-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific). To facilitate the occurrence of phage outside the carbon support, two rounds of sample application and blotting were used as described by Snijder *et al*[27]. Alternatively, the sample was applied only once onto the abovementioned grids, which were coated with a 3 nm thick layer of carbon. The samples were analysed by Tecnai F20 TEM (Thermo Fisher Scientific) operated at 200 kV.

Host range of *S. epidermidis* bacteriophages. The host range of *S. epidermidis* phage ΦTÜB and phage PH15 was determined with 35 *S. epidermidis* strains. To determine the bacteriophage susceptibility of *S. epidermidis* strains and other bacterial species, we applied 10 µl of phage lysate in 10-fold, 100-fold, and 1000-fold dilutions of the routine test dilution (RTD). RTD is a phage suspension at a concentration that

produces semi-confluent lysis on the reference strain after the application of 10 µl of phage suspension.

Phage isolation, propagation, and purification. The phage ΦTÜB was induced from a clinical *S. epidermidis* strain from infected teeth of a patient from the Hospital of Tübingen University. Briefly, 1 µg/ml mitomycin C was added to an overnight bacterial culture diluted to an OD_{600} value of 0.4. The culture was incubated at 30°C for 4 hours. The culture was filtered to obtain a sterile supernatant. The laboratory strain *S. epidermidis* 1457 was used for the separation of a single phage, detection of phage existence and propagation of the phage. ΦTÜB was propagated was under 37°C without shaking for 4 hours with an appropriate concentration of *S. epidermidis* 1457 according to the starting phage titer. The phage purification process was performed with centrifugation and filtering. The phage was concentrated by ultracentrifugation (Beckman Coulter).

Genome extraction of bacteria and phage. A Phage DNA Isolation Kit (Norgen Biotek) was used to extract the genome of ΦTÜB. In total, 1×10^{12} PFU of ΦTÜB were used. The genomic DNA of *S. epidermidis* E72 was extracted and quantified using the DNeasy Blood & Tissue Kit (Qiagen).

Genome sequencing and bioinformatic analysis of phage. The phage genome was assembled using Unicycler ver. 0.4.5.0 28 from a random subsample of 100,000 reads (RNG seed = 4), which was obtained by seqtk_sample ver. 1.2.0 (The Galaxy Project ver. 18.09; usegalaxy.org) from the raw sequencing data trimmed by Trimmomatic ver. 0.36.5[29]. Then, the genome assembly was verified, and the type of phage genome was predicted by alignment with the full dataset of trimmed reads in CLC Genomic Workbench 3.6.5 (QIAGEN Bioinformatics, Denmark) with manual inspection. The assembly parameters were set as follows: minimum length fraction = 0.6, similarity = 0.9, and ignoring non-specific matches. In total, 97.8% of 21,589,832 paired-end reads matched the assembled genome. The coverage ranged between 7,301 and 122,882, with an average of 56,023. Genomes were annotated using RAST[30], BLAST (http://blast.ncbi.nlm.nih.gov/), and InterPro (http://www.ebi.ac.uk/interpro). Protein and nucleotide alignments were calculated by EMBOSS Needle

(https://www.ebi.ac.uk/Tools). The genome comparison figure was created using EasyFig 2.2.2. (Sullivan *et al.,* 2011). Transmembrane helices in proteins were predicted using TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/).

Bacterial genome sequencing. Library preparation was conducted in accordance with the Nextera XT DNA Library Prep Guide (Illumina) or NEBNext Ultra library prep Kit for Illumina (New England Biolabs) after the DNA fragments greater than 300 bp were sheared on a Bioruptor Pico instrument (Diagenode). The libraries were sequenced on a MiSeq platform (Illumina) with 2 × 251 bp using a MiSeq Reagent Kit v2 or on a NextSeq Platform with 2x150 bp using a NextSeq 500/550 v2 Kit. Velvet or SPAdes was used to generate de novo assemblies.

Adsorption kinetics. Adsorption kinetics were determined as described previously[31]. Briefly, the adsorption was analysed using an MOI of 0.1, and the adsorption rate (%) was calculated by determining the number of unbound phage particles in the supernatant and subtracting it from the total number of input PFU as a ratio to the total number of input PFU. The adsorption rate was estimated 2 min after phage infection.

pH and temperature test. For storage tests in media with different pH values, 1 M NaOH and 1 M HCl were used to adjust the pH values of phage medium (TSB with 4 mM $CaCl_2$). ΦTÜB in 1 ml of medium was shaken at 300 rpm. For the storage test at different temperatures, ΦTÜB in 1 ml of medium was shaken at 300 rpm in a Thermoblock at different temperatures. After 24 h of incubation, the titer of ΦTÜB was tested. Both assays used phages with the same starting concentration of approximately 1×10^8 PFU/ml.

Electroporation. To test the electroporation efficiency of *S. epidermidis* strains, the pRB474 plasmid was used as the testing plasmid. The method of preparation of electrocompetent cells was performed according to Monk *et al*[32]. Briefly, overnight cultures of *S. epidermidis* strains were diluted in 10 ml of TSB to OD_{600}=0.5 and reincubated for 30 min. The bacterial cells were harvested by centrifugation at 4°C. The pellets were resuspended in an equal volume of sterile ice-cold water. These washing steps were repeated one more time. The cells were then washed in 1/10, 1/25,

and 1/200 volume of sterile ice-cold with 10% (W/Vol) glycerol. Aliquots of 50 µl were frozen at -80°C. The cells for electroporation were first thawed on ice for 5 min and then kept at room temperature for another 5 min. The cells were centrifuged and resuspended in 50 µl of filter-sterilized 500 mM sucrose with 10% glycerol. Five micrograms of plasmid in water was added to 50 µl of competent cells and transferred into an electroporation cuvette (Cell Project) with a 1 mm gap. A multiporator (Eppendorf) at a voltage of 1000 was used for electroporation. One millilitre of prewarmed fresh TSB was added to the cells and incubated at 37°C with shaking at 160 rpm. After 1 h of incubation, the cells were plated on TSB blood agar with 10 µg ml^{-1} of Cm and incubated at 37°C for 24 h.

Phage transduction. To test the transduction efficiency in *S. epidermidis*, the pBTn plasmid[33] was used. Plasmid transduction in *S. epidermidis* strains was performed using ΦTÜB as the transducing phage and 1457 as the donor strain or Φ187 as the transducing phage and PS187 as the donor strain according to the method by Winstel *et al*[17]. Briefly, 1457 or PS187 bearing the testing plasmid pBTn was infected with ΦTÜB or Φ187, and the lysate was used to infect recipient *S. epidermidis* strains. Bacteria were grown overnight and resuspended in phage buffer (100 mM $MgSO_4$, 100 mM $CaCl_2$, 1 M Tris–HCl, pH 7.8, 0.59% NaCl, 0.1% gelatine); the bacteria were mixed with 100 µl of lysate (~1×10^9 PFU ml^{-1}) and incubated at 37°C for 10 min. The mixture was then plated on TSB agar plates containing 10 µg ml^{-1} chloramphenicol.

Result and discussion

The novel phage ΦTÜB from a clinical *S. epidermidis* strain. A total of 114 clinical *S. epidermidis* strains were collected from the Medical Microbiology Department at the Hospital of Tübingen University. For three of these 114 *S. epidermidis* strains, phages could be induced by mitomycin C. Nine clinical *S. epidermidis* strains together with one laboratory *S. epidermidis* strain, 1457, were used for the lysis assay for phage detection. The phage lysate of strain E72, a clinical *S. epidermidis* strain from infected teeth of a patient, was found to have the ability to infect more *S. epidermidis* strains than the other two phage lysates induced. In addition, this lysate could also infect the easily electroporated laboratory strain *S. epidermidis* 1457. A single phage plaque was isolated, propagated and tested; the phage was able to infect both 1457 and a relatively large number of clinical *S. epidermidis* strains. This phage was named ΦTÜB.

Analysis of the *S. epidermidis* E72 genome. Analysis of the E72 genome showed that the prophage of ΦTÜB was located immediately downstream of the tRNA-Ser gene.

Analysis of the ΦTÜB genome. The genome of ΦTÜB was compared those of *S. aureus* Φ11 and Φ187, the first published *S. epidermidis* genome of ΦPH15, and the genomes of other *S. epidermidis Siphoviridae* from the NCBI database (Figure 1). All genomes were modularly organized and consisted of DNA packaging, DNA metabolism, head and tail morphogenesis, lysogenic conversion, host cell lysis, and lysogeny modules. The genome of ΦTÜB consists of 44,592 bp of dsDNA with a GC content of 34.5%. It probably belongs to headful packaging phages with circularly permutated genomes with redundant ends, as no physical termini were obvious based on a sequencing read alignment according to Garneau *et al*[34]. The ΦTÜB genome contains 71 predicted CDS (coding sequences) ranging from 105 bp (*gp_5* encoding a hypothetical protein) to 1034 bp (*gp_17* encoding a tape measure protein), of which 45 encode proteins with known or predicted functions. A non-coding intron was

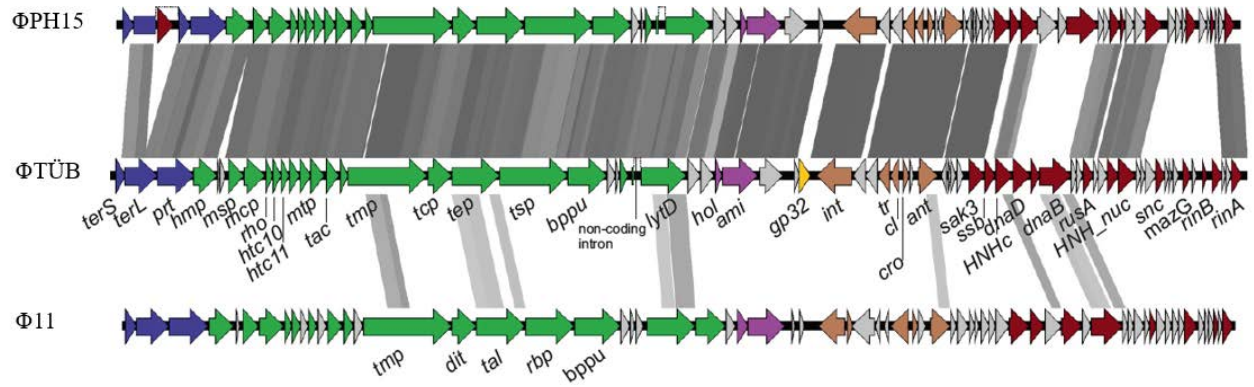
predicted in the cell wall hydrolase gene *gp_25*. Contrary to phage ΦPH15 (NC_008723), which shows 86.3% identity to ΦTÜB, ΦTÜB does not contain a coding intron in the gene for the terminase large subunit. These two phages differ mostly in the DNA metabolism module. The structural modules, such as head and tail morphogenesis and DNA packaging, of most other *S. epidermidis* phages, including phiCNPH82 (NC_008722), phiCNPx (KU598975), vB_SepiS-phiIPLA5 (NC_018281), vB_SepiS-phiIPLA7 (NC_018284), and phiIME1348_01 (KY653120), showed high similarities with the ΦTÜB genome.

Gene *gp_32* encoding a putative membrane-associated protein is localized in the region between the lytic module and the integrase gene. As this represents the correct terminus of prophage, the gene could be obtained by aberrant prophage excision. Gene *gp_32* is identical to the DUF2335 domain-containing protein of *S. epidermidis* (WP_103433786). Its first inside region is composed of amino acids in pos. 1 - 95, the first transmembrane helix in pos. 96 - 118, the outside region in pos. 119 - 121, the second transmembrane helix in pos. 122 - 141, and the second inside region in pos. 142 - 149 (computed using http://www.cbs.dtu.dk/services/TMHMM/). Due to its transmembrane character, it is possible that gene *gp_32* is a part of some system connected with bacterial virulence, but the exact function is unknown. None of the similar phages encode this protein. ΦPH15 contains almost identical sequences (99% identity) in the same region but with one-nucleotide insertion causing a premature stop codon after the third triplet of the corresponding reading frame.

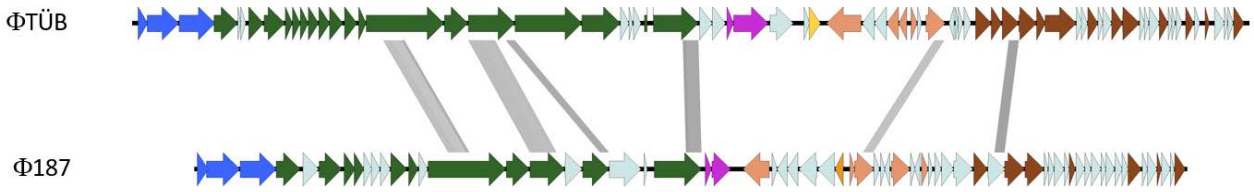
The genes localized between the tape measure protein gene and the lysis module were previously shown to be essential for Φ11 (NC_004615) host recognition and infectivity[35,36]. These genes include *gp43*, *gp44*, *gp45* and *gp54*, encoding the distal tail protein (Dit), tail associated lysin (Tal), receptor-binding protein (RBP) and upper baseplate protein (BppU), respectively. Some similarities were found between the products of the host recognition genes of Φ11 and the corresponding gene products of ΦTÜB. The Φ11 protein Dit shows 30.1% identity and 45.9% similarity, Tal shows

51.6% identity and 69.6% similarity, RBP shows 13.1% identity and 21.3% similarity, and BppU shows 20.7% identity and 37.1% similarity to the proteins encoded by ΦTÜB genes *gp_18 (tcp)*, *gp_19 (tep)*, *gp_20 (tsp)*, and *gp_21 (bppu)*, respectively. Φ187 (NC_007047) showed even less similarity to ΦTÜB. Genomes of *S. epidermidis* phages phiStB20-like (NC_028821), phiHOB14.1.R1 (CP018841), phivB_SepS_SEP9 (NC_023582), and phiSPbeta-like (NC_029119) shared almost no similarity to ΦTÜB.

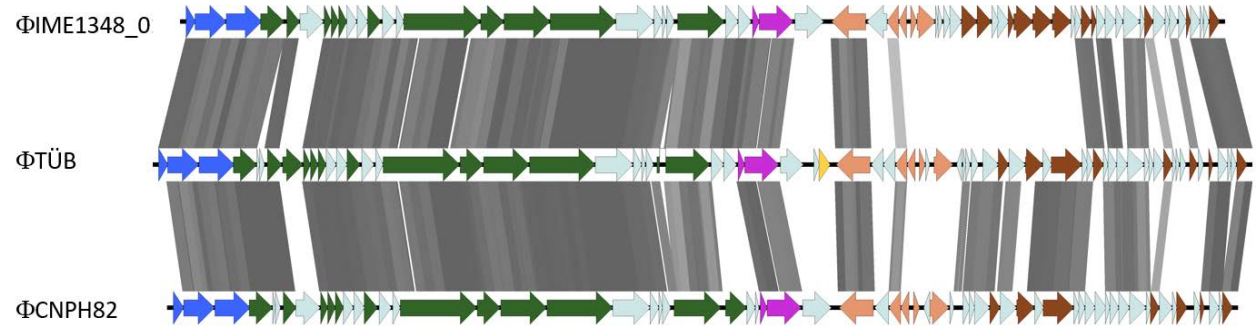
a



b



c



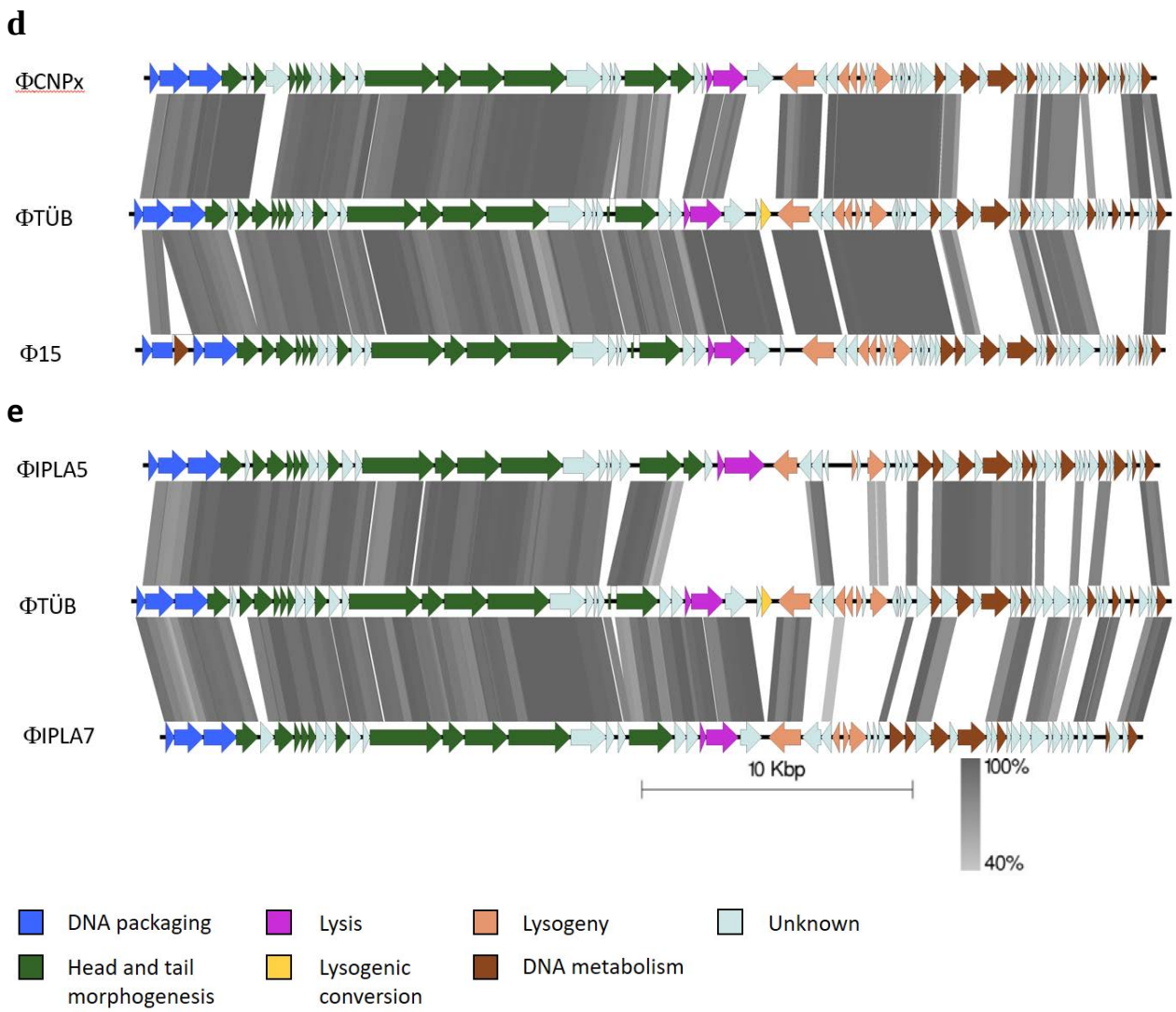


Figure 1. Comparison of *S. epidermidis* ΦTÜB to various bacteriophage genomes. Predicted genes (listed in Table S1) are coloured with respect to genome modules and are represented by full arrows. Parts of genes divided by introns are joined by the dashed line. Shaded boxes express the sequence identity levels between genomes computed using tblastx and ranging from 40% to 100%.

***S. epidermidis* phage ΦTÜB belongs to the *Siphoviridae* family, the infection range of which is usually narrow to one species**. Morphological analysis by transmission electron microscopy (TEM) with negative staining and cryo-conditions (cryo-EM) showed that ΦTÜB has an icosahedral head (B1 morphology) with a flexible, non-contractile tail ending with a baseplate. The diameter of the head is 65 nm, the length of the tail is 155 nm, and the length of the baseplate is 27 nm (Figure 2). Therefore, it is considered a phage of the family *Siphoviridae*. The adsorption kinetics of *S. epidermidis* phage ΦTÜB and *S. aureus* phage Φ11 were determined on *S. epidermidis* strain 27, *S. epidermidis* strain 15 and *S. aureus* strain ISP8 (Figure 3). ΦTÜB efficiently adsorbed onto the *S. epidermidis* strains and did not adsorb onto the *S. aureus* strain. However, Φ11 showed similarity with ΦTÜB in host recognition genes and adsorbed efficiently onto the *S. aureus* strain; furthermore, reversible adsorption was observed in the case of both *S. epidermidis* strains. The host range of ΦTÜB was found to have a narrow host range only to *S. epidermidis* when determined on various bacterial species (Table 1a,1b). A set of 35 *S. epidermidis* strains was also used for the detection of the sensitivity of ΦTÜB to infect *S. epidermidis*. ΦTÜB could infect more strains than *S. epidermidis* phage ΦPH15 (Table 1c).

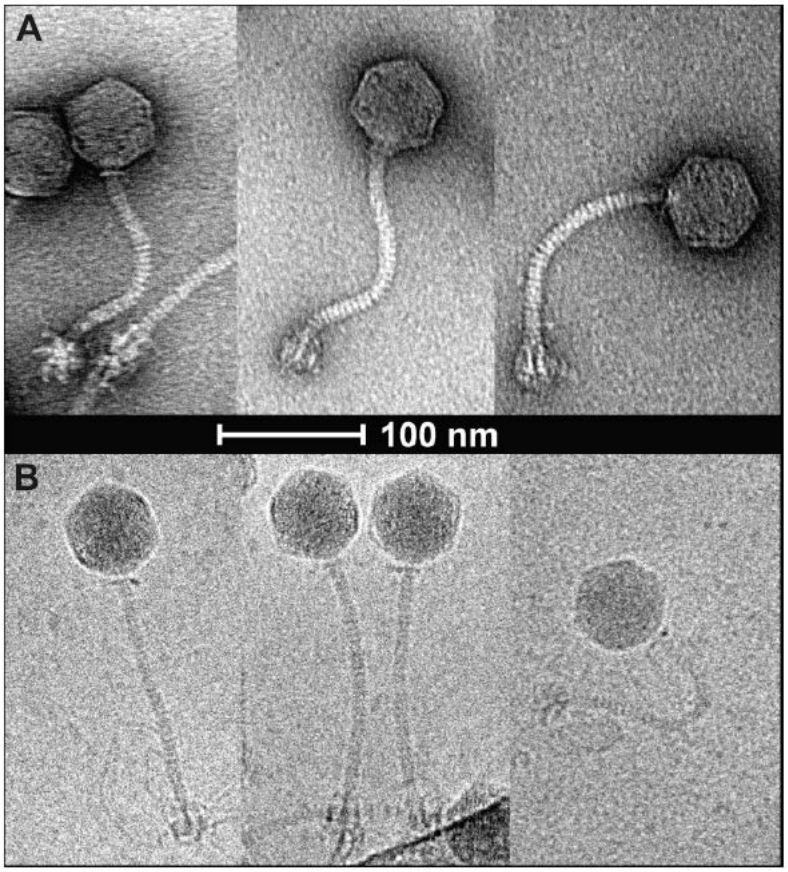


Figure 2. TEM images of the phage ΦTÜB. (A) The negatively stained particles with different orientations of the baseplate. The baseplate contains six receptor-binding structures. **(B)** The cryo-preserved native particles. The collar complex of the phage is noticeable. The baseplate tends to adsorb on loose DNA strands, other phages or carbon.

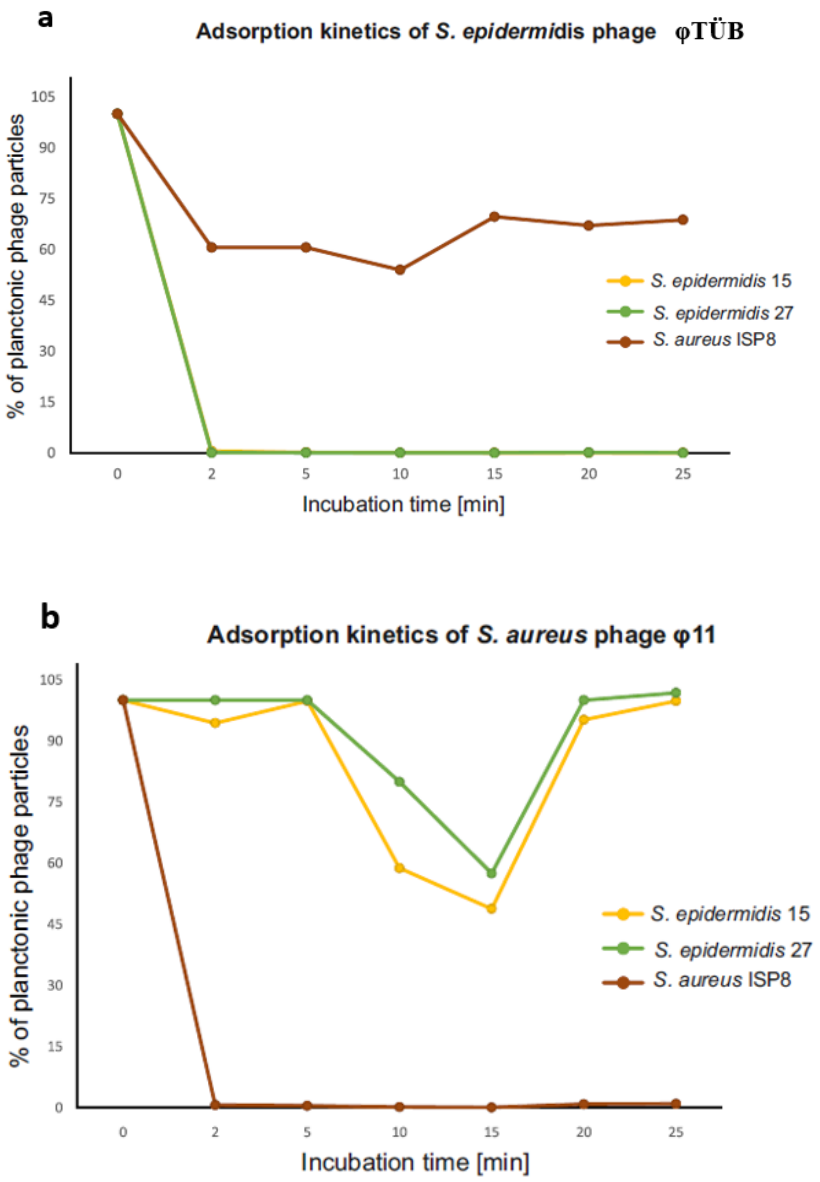


Figure 3. Adsorption kinetics of *S. epidermidis* phage ΦTÜB and *S. aureus* phage Φ11. The adsorption rate (%) was calculated by determining the number of unbound phage particles in the supernatant and subtracting it from the total number of input PFU. The results were expressed as a percentage of the initial phage count.

Table 1. Host range of *S. epidermidis* bacteriophages ΦTÜB. (a). Sensitivity test on different bacterial species; (b). Sensitivity test on *S. epidermidis*; (c). Comparison of the sensitivity of ΦPH15 and ΦTÜB to methicillin-resistant and methicillin-sensitive *S. epidermidis*.

a

Species	Strain	Sensitivity
S. aureus	RN4220	-
	PS187	-
	770wt	-
	Newman	-
	ST 398	-
	T166-1	-
	ATCC6538	-
	ATCC33591	-
	SA113	-
S. carnosus	TM300	-
S. caprae	BK14568/12	-
S. capitis	1125	-
S. cohnii	1124	-
S. hominis	1126	-
S. pasteuri	1127	-
S. simulans	1129	-
S. warneri	IVK 51	-
S. equorum	LTH5015	-
S. lugdunensis	1005 wt	-
E. faecalis	ATCC 29212	-
E. faecalis	583	-
E. faecium	ST4144	-
E. faecium	BK4705	-
Bacillus subtilis	BEST195	-
Listeria grayi	ATCC25401	-
Listeria nonocytogenes	ATCC19118	-
S. saprophyticus	NT219	-
S. pyogenes	BK2192	-

b

Species	Strain	Sensitivity
S. epidermidis	1457	+
	RP62A	-
	ATCC12228	-
	Tu3298	-
	IVK83	-
S. epidermidis	E63	-
(from infection sites)	E64	+
	E65	+
	E66	+
	E67	-
	E68	-
	E69	-
	E70	-
	E71	-
S. epidermidis	B3-16	-
(from nasal colonization)	B3-18	+
	B3-26	-
	B4-1	-
	B4-10	+
	B4-12	+
	B4-18	-
	B4-20	-
	B4-21	-
	B4-24	-
	B4-25	+
	B4-39	-
	B4-40	+
	B4-41	-

c

***S. epidermidis* strains**	**Sensivity**	
	ΦPH15	**ΦTÜB**
15	S	S
27	S	S
48	R	R
456	R	S
459	S	S
471	R	S
B1	S	S
A6C	S	S
AqC	S	S
62-A	R	R
O-47	R	R
CCM 50	R	R
CCM 4187	R	R
CCM 2343	R	R
CCM 4418	R	R
CCM 7844	R	S
CCM 7221	R	R
CCM 2124	R	S
CCM 4505	R	S
CCM 3298	R	R
15 methicillin-resistant clinical *S. epidermidis* strains	none of 15 strains susceptible	one of 15 strains susceptible

R - strain is resistant to bacteriophage; S - strain is sensitive to bacteriophage

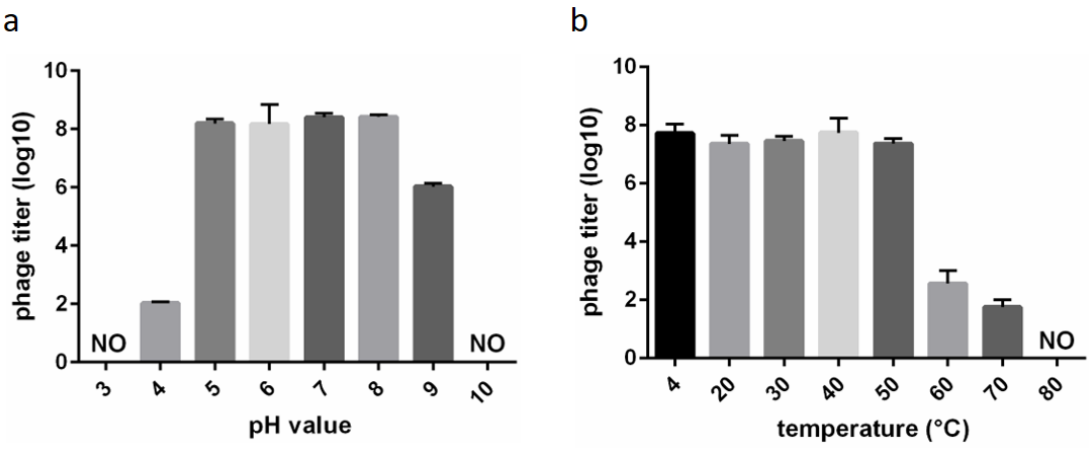


Figure 4. Stable conditions for ΦTÜB. (a) Sensitivity to conditions with different pH values; (b) sensitivity to conditions with different temperatures.

ΦTÜB is stable in neutral medium and at mild temperature. ΦTÜB could be stable in TSB medium at pH values between 5.0 and 8.0 and temperatures from 4-50°C (Figure 4a,4b).

ΦTÜB can transduce plasmid DNA with high efficiency to a large percentage of *S. epidermidis* strains that cannot be transformed via electroporation. The transformation of the plasmid pBTn to most of the 37 clinical *S. epidermidis* strains (26/37, 70.3%) via the electroporation method failed. The plasmid entrance to the cells is the first step of the molecular study of bacteria. As an alternative strategy, the transduction of plasmid DNA by phages is promising. Winstel *et al*[17] established a phage transduction method for CoNS strains using Φ187, a phage that can transduce both *S. aureus* and CoNS with high efficiency. The Φ187 transduction method could not apply to all *S. epidermidis* strains. Therefore, a novel phage to cover more *S. epidermidis* is needed. First, a host strain with similar DNA methylation systems as the target *S. epidermidis* strains should be chosen. A strain from the same species is preferred. The laboratory strain *S. epidermidis* 1457 was chosen as a strain more easily transformed by electroporation. ΦTÜB was first propagated until it reached a

titer higher than 1×10^9 PFU for each ml of phage lysate. The freshly propagated ΦTÜB was propagated again once with host strain 1457, which was transformed with plasmid pBTn to create a phage lysate with pBTn. Our results showed that this transduction method with the novel phage ΦTÜB (Figure 5) facilitated high efficiency plasmid entrance into the bacteria cells with frequencies up to 10^4 transductants per ml of phage lysate to a large percentage of *S. epidermidis* strains (31/37, 83.8%), even (19/22, 86.4%) to those strains that could not be transformed by electroporation or be transduced by Φ187 (Figure 6). ΦTÜB is a promising alternative tool for the phage transduction of plasmids for *S. epidermidis*.

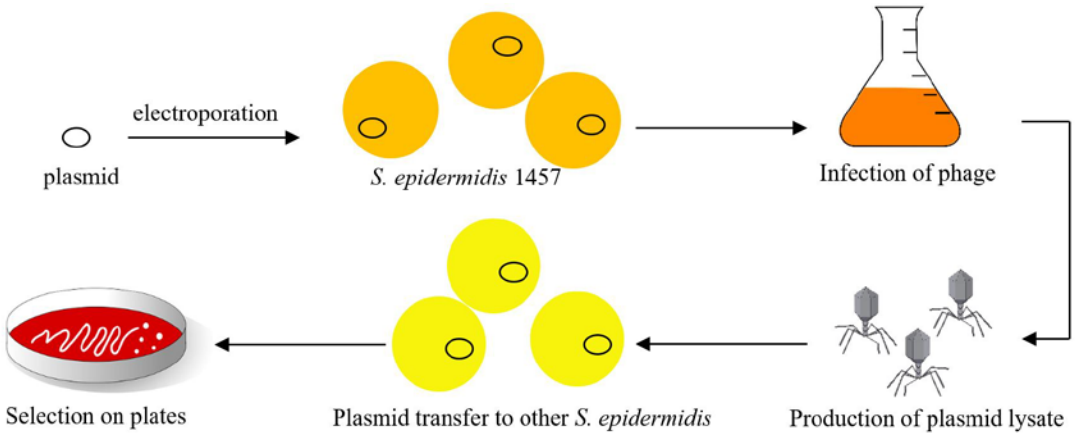


Figure 5. Schematic of plasmid transduction by ΦTÜB.

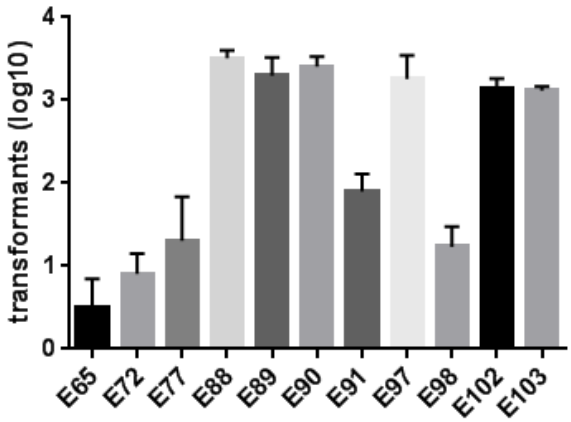


Figure 6. Transduction of plasmids into *S. epidermidis* using ΦTÜB. (a) Plasmid transfer mediated by ΦTÜB in selected clinical *S. epidermidis* strains that can be infected by ΦTÜB but cannot be transformed by electroporation or transduction by Φ187 (part of the results are shown).

Conclusion

In this study, we induced a novel phage, ΦTÜB, from a clinical *S. epidermidis* strain. We investigated the genome characteristics of ΦTÜB and demonstrated that ΦTÜB is highly similar to the first reported *S. epidermidis* phage ΦPH15 and most other *S. epidermidis Siphoviridae* but not to *S. aureus* phages Φ11 or Φ187. ΦTÜB has a narrow host range to *S. epidermidis* but a broader host range to strains. Most importantly, ΦTÜB showed the potential to transduce plasmids into a high percentage of *S. epidermidis* with high efficiency. We can use ΦTÜB to transform plasmids as an alternative to electroporation and Φ187 transduction. Further study focusing on the bacteria receptor for ΦTÜB is expected to provide better insight into the mechanism and role of phages in the pathogenesis and evolution of *S. epidermidis*.

Table S1. Genes and encoded proteins of ΦTÜB and the closest phage protein matches.

Gene	Strand	Region	Product	Size (aa)	Best GenBank hit (blastp)		Identity (%)
					Protein Accession No.	Reference phage (Accesion No.)	
gp_1 (terS)	+	239..619	terminase small subunit	126	ARM68050	IME1348_01 (KY653120)	99
gp_2 (terL)	+	603..1868	terminase large subunit	421	YP_950600	CNPH82 (NC_008722)	98
gp_3 (prt)	+	1874..3310	portal protein	472	ARM68052	IME1348_01 (KY653120)	99
gp_4 (hmp)	+	3267..4220	head morphogenesis protein	317	YP_950602	CNPH82 (NC_008722)	100
gp_5	+	4223..4327	hypothetical protein	34	YP_950603	CNPH82 (NC_008722)	100
gp_6	+	4328..4534	hypothetical protein	68	YP_950604	CNPH82 (NC_008722)	100
gp_7 (msp)	+	4649..5245	scaffolding protein	198	YP_006560947	IPLA5 (NC_018281)	95
gp_8 (mcp)	+	5263..6093	major capsid protein	276	YP_006561168	IPLA7 (NC_018284)	99
gp_9 (rho)	+	6110..6400	Rho termination factor	96	YP_006561169	IPLA7 (NC_018284)	100
gp_10 (htc10)	+	6400..6714	head-tail adapter protein	104	YP_950608	CNPH82 (NC_008722)	100
gp_11 (htc11)	+	6707..7036	head-tail adapter protein	109	YP_950609	CNPH82 (NC_008722)	99
gp_12	+	7029..7442	HK97 gp10 family phage protein	137	YP_950610	CNPH82 (NC_008722)	100
gp_13	+	7455..7892	tail completion protein	145	YP_950611	CNPH82 (NC_008722)	99
gp_14 (mtp)	+	7879..8418	tail protein	179	YP_006560954	IPLA5 (NC_018281)	98
gp_15 (tac)	+	8481..8975	tail assembly chaperone	164	YP_006560955	IPLA5 (NC_018281)	100
gp_16	+	9038..9340	tail protein	100	YP_950677	PH15 (NC_008723)	99
gp_17 (tmp)	+	9343..12447	tape measure protein	1034	YP_950615	CNPH82 (NC_008722)	99

gp_18 (tcp)	+	12463..13401	siphovirus-type tail component	312	YP_006560958	IPLA5 (NC_018281)	99
gp_19 (tep)	+	13415..15271	putative tail endopeptidase	618	ARM68066	IME1348_01 (KY653120)	99
gp_20 (tsp)	+	15286..17952	putative glycoside hydrolase	888	YP_950618	CNPH82 (NC_008722)	99
gp_21 (bppu)	+	17952..19484	baseplate upper protein	510	YP_950619	CNPH82 (NC_008722)	99
gp_22	+	19489..19827	hypothetical protein	112	YP_950620	CNPH82 (NC_008722)	100
gp_23	+	19829..19969	XkdX family protein	46	YP_950621	CNPH82 (NC_008722)	100
gp_24	+	20006..20305	membrane-associated protein	99	YP_950685	PH15 (NC_008723)	79[1]
gp_25 (lytD)	+	20777..22591	cell wall hydrolase	633	YP_950686	PH15 (NC_008723)	97
gp_26	+	22644..23153	hypothetical protein	169	YP_950687	PH15 (NC_008723)	99
gp_27	+	23153..23680	hypothetical protein	175	YP_006561186	IPLA7 (NC_018284)	99
gp_28 (hol)	+	23736..24002	holin	88	YP_950689	PH15 (NC_008723)	100
gp_29 (ami)	+	24002..25384	N-acetylmuramoyl-L-alanine amidase	460	YP_950690	PH15 (NC_008723)	100
gp_30	+	25462..26388	hypothetical protein	308	YP_006561189	IPLA7 (NC_018284)	100
gp_31	+	26817..27032	hypothetical protein	71	YP_950692	PH15 (NC_008723)	100
gp_32	+	27022..27471	membrane-associated protein	149	YP_009219652	IME-SA4 (NC_029025)	47[1]
gp_33 (int)	-	27712..29088	integrase	458	YP_950693	PH15 (NC_008723)	99
gp_34	-	29142..29624	hypothetical protein	160	YP_950694	PH15 (NC_008723)	100
gp_35	-	29626..30102	hypothetical protein	158	YP_950695	PH15 (NC_008723)	100
gp_36 (tr)	-	30121..30582	putative transcriptional regulator	153	YP_950696	PH15 (NC_008723)	100
gp_37 (cI)	-	30594..30917	putative cI-like repressor	107	YP_950697	PH15 (NC_008723)	100
gp_38 (cro)	+	31081..31329	putative Cro-like repressor	82	YP_950698	PH15 (NC_008723)	100
gp_39	+	31342..31503	hypothetical protein	53	YP_950699	PH15 (NC_008723)	100
gp_40 (ant)	+	31683..32450	antirepressor protein	255	YP_950701	PH15 (NC_008723)	100
gp_41	-	32603..32809	hypothetical protein	68	YP_950703	PH15 (NC_008723)	100
gp_42	+	32856..32969	hypothetical protein	37	YP_009302060	CNPx (NC_031241)	100
gp_43	+	32966..33142	hypothetical protein	58	YP_950705	PH15 (NC_008723)	100
gp_44	+	33202..33477	hypothetical protein	91	YP_950706	PH15 (NC_008723)	100
gp_45 (sak3)	+	33666..34292	single-strand annealing protein SAK3	208	YP_950708	PH15 (NC_008723)	100
gp_46 (ssb)	+	34285..34707	single-stranded DNA-binding protein	140	YP_006561200	IPLA7 (NC_018284)	100
gp_47 (HNHc)	+	34721..35395	putative HNHc nuclease	224	YP_006560984	IPLA5 (NC_018281)	100
gp_48 (dnaD)	+	35392..36093	DNA replication protein DnaD	233	YP_006560985	IPLA5 (NC_018281)	100
gp_49	+	36099..36452	putative replisome organizer	117	ASN69727	10AX_1 (MF417895)	100
gp_50 (dnaB)	+	36439..37689	replicative DNA helicase DnaB	416	YP_950645	CNPH82 (NC_008722)	100
gp_51	+	37686..37907	hypothetical protein	73	ASN71264	10F_5 (MF417921)	93
gp_52	+	37885..38130	hypothetical protein	81	YP_006561206	IPLA7 (NC_018284)	99
gp_53 (rusA)	+	38139..38546	Holliday junction resolvase RusA	135	YP_950716	PH15 (NC_008723)	100
gp_54	+	38547..38738	hypothetical protein	63	ASN70066	9S_1 (MF417901)	90

gp_55	+	38739..39098	phiPVL ORF050-like protein	119	ASN70067	9S_1 (MF417901)	97
gp_56	+	39095..39529	putative nucleotide kinase	144	YP_950719	PH15 (NC_008723)	99
gp_57 (HNH_nuc)	+	39532..40212	endonuclease	226	ARM68038	IME1348_01 (KY653120)	97
gp_58	+	40209..40406	hypothetical protein	65	ASN70071	9S_1 (MF417901)	88
gp_59	+	40393..40572	hypothetical protein	59	YP_006561212	IPLA7 (NC_018284)	95
gp_60	+	40591..40989	hypothetical protein	132	APC43045	StAP1 (KX532239)	31[1]
gp_61 (snc)	+	40993..41346	nuclease	117	ASN70073	9S_1 (MF417901)	92
gp_62	+	41351..41503	hypothetical protein	50	YP_006561001	IPLA5 (NC_018281)	100
gp_63	+	41493..41627	hypothetical protein	44	not annotated	IPLA7 (NC_018284)	99
gp_64	+	41659..41895	hypothetical protein	78	YP_006561216	IPLA7 (NC_018284)	100
gp_65 (mazG)	+	42042..42365	putative NTP pyrophosphohydrolase	107	YP_006561217	IPLA7 (NC_018284)	99
gp_66	+	42402..42605	hypothetical protein	67	YP_009214584	IPLA-C1C (NC_028962)	96
gp_67 (rinB)	+	42824..42994	putative regulator RinB	56	YP_950727	PH15 (NC_008723)	96
gp_68	+	43200..43598	membrane-associated protein	132	YP_950661	CNPH82 (NC_008722)	100
gp_69	+	43586..43726	hypothetical protein	46	YP_950662	CNPH82 (NC_008722)	100
gp_70	+	43730..43948	hypothetical protein	72	YP_006561220	IPLA7 (NC_018284)	100
gp_71 (rinA)	+	43966..44382	RinA family transcriptional regulator	138	YP_950730	PH15 (NC_008723)	100

References

1 Oh, J. *et al.* Biogeography and individuality shape function in the human skin metagenome. *Nature* **514**, 59-64, doi:10.1038/nature13786 (2014).

2 Grice, E. A. *et al.* Topographical and temporal diversity of the human skin microbiome. *Science* **324**, 1190-1192, doi:10.1126/science.1171700 (2009).

3 Oh, J., Conlan, S., Polley, E. C., Segre, J. A. & Kong, H. H. Shifts in human skin and nares microbiota of healthy children and adults. *Genome Med* **4**, 77, doi:10.1186/gm378 (2012).

4 Otto, M. Staphylococcus epidermidis--the 'accidental' pathogen. *Nat Rev Microbiol* **7**, 555-567, doi:10.1038/nrmicro2182 (2009).

5 Queck, S. Y. *et al.* Mobile genetic element-encoded cytolysin connects virulence to methicillin resistance in MRSA. *PLoS Pathog* **5**, e1000533, doi:10.1371/journal.ppat.1000533 (2009).

6 Li, M., Wang, X., Gao, Q. & Lu, Y. Molecular characterization of Staphylococcus epidermidis strains isolated from a teaching hospital in Shanghai, China. *J Med Microbiol* **58**, 456-461, doi:10.1099/jmm.0.007567-0 (2009).

7 Rohde, H. *et al.* Polysaccharide intercellular adhesin or protein factors in biofilm accumulation of Staphylococcus epidermidis and Staphylococcus aureus isolated from prosthetic hip and knee joint infections. *Biomaterials* **28**, 1711-1720, doi:10.1016/j.biomaterials.2006.11.046 (2007).

8 Delgado, S. *et al.* Staphylococcus epidermidis strains isolated from breast milk of women suffering infectious mastitis: potential virulence traits and resistance to antibiotics. *BMC Microbiol* **9**, 82, doi:10.1186/1471-2180-9-82 (2009).

9 Deb, R. *et al.* Trends in diagnosis and control of bovine mastitis: a review. *Pak J Biol Sci* **16**, 1653-1661 (2013).

10 Otto, M. Coagulase-negative staphylococci as reservoirs of genes facilitating MRSA infection: Staphylococcal commensal species such as Staphylococcus epidermidis are being recognized as important sources of genes promoting MRSA colonization and virulence. *Bioessays* **35**, 4-11, doi:10.1002/bies.201200112 (2013).

11 Du, X. *et al.* Molecular analysis of Staphylococcus epidermidis strains isolated from community and hospital environments in China. *PLoS One* **8**, e62742, doi:10.1371/journal.pone.0062742 (2013).

12 Witte, W. *et al.* Emergence and spread of antibiotic-resistant Gram-positive bacterial pathogens. *Int J Med Microbiol* **298**, 365-377, doi:10.1016/j.ijmm.2007.10.005 (2008).

13 Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843-1845, doi:10.1126/science.1165771 (2008).

14 Costa, S. K., Donegan, N. P., Corvaglia, A. R., Francois, P. & Cheung, A. L. Bypassing the Restriction System To Improve Transformation of Staphylococcus epidermidis. *J Bacteriol* **199**, doi:10.1128/JB.00271-17 (2017).

15 Gotz, F., Ahrne, S. & Lindberg, M. Plasmid transfer and genetic recombination by protoplast fusion in staphylococci. *J Bacteriol* **145**, 74-81 (1981).

16 Robinson, J. M., Hardman, J. K. & Sloan, G. L. Relationship between lysostaphin endopeptidase production and cell wall composition in Staphylococcus staphylolyticus. *J Bacteriol* **137**, 1158-1164 (1979).

17 Winstel, V., Kuhner, P., Krismer, B., Peschel, A. & Rohde, H. Transfer of plasmid DNA to clinical coagulase-negative staphylococcal pathogens by using a unique bacteriophage. *Appl Environ Microbiol* **81**, 2481-2488, doi:10.1128/AEM.04190-14 (2015).

18 Maslanova, I., Stribna, S., Doskar, J. & Pantucek, R. Efficient plasmid transduction to Staphylococcus aureus strains insensitive to the lytic action of

transducing phage. *Fems Microbiol Lett* **363**, doi:10.1093/femsle/fnw211 (2016).

19 Varga, M., Pantucek, R., Ruzickova, V. & Doskar, J. Molecular characterization of a new efficiently transducing bacteriophage identified in meticillin-resistant Staphylococcus aureus. *J Gen Virol* **97**, 258-268, doi:10.1099/jgv.0.000329 (2016).

20 Dowell, C. E. & Rosenblum, E. D. Serology and Transduction in Staphylococcal Phage. *J Bacteriol* **84**, 1071-1075 (1962).

21 Jefferson, S. J. & Parisi, J. T. Bacteriophage typing of coagulase-negative staphylococci. *J Clin Microbiol* **10**, 396-397 (1979).

22 Daniel, A., Bonnen, P. E. & Fischetti, V. A. First complete genome sequence of two Staphylococcus epidermidis bacteriophages. *J Bacteriol* **189**, 2086-2100, doi:10.1128/JB.01637-06 (2007).

23 Gutierrez, D., Martinez, B., Rodriguez, A. & Garcia, P. Genomic characterization of two Staphylococcus epidermidis bacteriophages with anti-biofilm potential. *BMC Genomics* **13**, 228, doi:10.1186/1471-2164-13-228 (2012).

24 Gutierrez, D., Martinez, B., Rodriguez, A. & Garcia, P. Isolation and characterization of bacteriophages infecting Staphylococcus epidermidis. *Curr Microbiol* **61**, 601-608, doi:10.1007/s00284-010-9659-5 (2010).

25 Melo, L. D. *et al.* Characterization of Staphylococcus epidermidis phage vB_SepS_SEP9 - a unique member of the Siphoviridae family. *Res Microbiol* **165**, 679-685, doi:10.1016/j.resmic.2014.09.012 (2014).

26 Melo, L. D. *et al.* Isolation and characterization of a new Staphylococcus epidermidis broad-spectrum bacteriophage. *J Gen Virol* **95**, 506-515, doi:10.1099/vir.0.060590-0 (2014).

27 Snijder, J. *et al.* Vitrification after multiple rounds of sample application and blotting improves particle density on cryo-electron microscopy grids. *J Struct Biol* **198**, 38-42, doi:10.1016/j.jsb.2017.02.008 (2017).

28 Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* **13**, e1005595, doi:10.1371/journal.pcbi.1005595 (2017).

29 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).

30 Brettin, T. *et al.* RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* **5**, 8365, doi:10.1038/srep08365 (2015).

31 Winstel, V. *et al.* Wall teichoic acid structure governs horizontal gene transfer between major bacterial pathogens. *Nat Commun* **4**, 2345, doi:10.1038/ncomms3345 (2013).

32 Monk, I. R., Shah, I. M., Xu, M., Tan, M. W. & Foster, T. J. Transforming the untransformable: application of direct transformation to manipulate genetically Staphylococcus aureus and Staphylococcus epidermidis. *MBio* **3**, doi:10.1128/mBio.00277-11 (2012).

33 Li, M. *et al.* Staphylococcus aureus mutant screen reveals interaction of the human antimicrobial peptide dermcidin with membrane phospholipids. *Antimicrob Agents Chemother* **53**, 4200-4210, doi:10.1128/AAC.00428-09 (2009).

34 Garneau, J. R., Depardieu, F., Fortier, L. C., Bikard, D. & Monot, M. PhageTerm: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. *Sci Rep* **7**, 8292, doi:10.1038/s41598-017-07910-5 (2017).

35 Li, X. *et al.* An essential role for the baseplate protein Gp45 in phage adsorption to Staphylococcus aureus. *Sci Rep* **6**, 26455, doi:10.1038/srep26455 (2016).

36 Koc, C. *et al.* Structural and enzymatic analysis of TarM glycosyltransferase from Staphylococcus aureus reveals an oligomeric protein specific for the glycosylation of wall teichoic acid. *J Biol Chem* **290**, 9874-9885, doi:10.1074/jbc.M114.619924 (2015).

Chapter 4

-

An accessory wall teichoic acid glycosyltransferase protects *Staphylococcus aureus* from the lytic activity of *Podoviridae*

Xuehua Li[1,2], David Gerlach[1,2], Xin Du[1,2], Jesper Larsen[3], Marc Stegger[3,4], Petra Kühner[1,2], Andreas Peschel[1,2], Guoqing Xia[1,2,5] & Volker Winstel[1,2]†.

1.Infection Biology, Interfaculty Institute of Microbiology and Infection Medicine, University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany. 2.German Center for Infection Research (DZIF), partner site Tübingen, 72076 Tübingen, Germany. 3.Microbiology and Infection Control, Statens Serum Institut, Artillerivej 5, 2300 Copenhagen, Denmark. 4.Pathogen Genomics Division, Translational Genomics Research Institute, 3051 WShamrell Blvd, Flagstaff, 86001 Arizona, USA. 5.Institute of Inflammation & Repair, The University of Manchester, Manchester, United Kingdom. †Present address: Institute for Medical Microbiology and Hospital Epidemiology, Hannover Medical School, Feodor-Lynen-Str. 7, 30625 Hannover, Germany. Correspondence and requests for materials should be addressed to V.W. (email: Winstel.Volker@mh-hannover.de)

Scientific Reports. 5, 17219; doi: 10.1038/srep17219 (2015).

Abstract

Many *Staphylococcus aureus* have lost a major genetic barrier against phage infection, termed clustered regularly interspaced palindromic repeats (CRISPR/cas). Hence, *S. aureus* strains frequently exchange genetic material via phage-mediated horizontal gene transfer events, but, in turn, are vulnerable in particular to lytic phages. Here, a novel strategy of *S. aureus* is described, which protects *S. aureus* against the lytic activity of *Podoviridae*, a unique family of staphylococcal lytic phages with short, non-contractile tails. Unlike most staphylococcal phages, *Podoviridae* require a precise wall teichoic acid (WTA) glycosylation pattern for infection. Notably, TarM-mediated WTA α-O-GlcNAcylation prevents infection of *Podoviridae* while TarS-mediated WTA β-O-GlcNAcylation is required for *S. aureus* susceptibility to podoviruses. Tracking the evolution of TarM revealed an ancient origin in other *staphylococci* and vertical inheritance during *S. aureus* evolution. However, certain phylogenetic branches have lost *tarM* during evolution, which rendered them podovirussusceptible. Accordingly, lack of *tarM* correlates with podovirus susceptibility and can be converted into a podovirus-resistant phenotype upon ectopic expression of *tarM* indicating that a "glycoswitch" of WTA O-GlcNAcylation can prevent the infection by certain staphylococcal phages. Since lytic staphylococcal phages are considered as anti-*S. aureus* agents, these data may help to establish valuable strategies for treatment of infections.

Introduction

Horizontal gene transfer (HGT) events are prerequisites for bacterial evolution. Bacteria, including many Gram-positive pathogens, employ different mechanisms for the exchange of genetic information. Major mechanisms include bacteriophage-(phage) mediated transduction, conjugation, and transformation[1,2]. These factors substantially contribute to bacterial evolution but vary in their impact depending on the bacterial species.

During evolution, many bacteria evolved various protective mechanisms that interfere with or impede HGT events. “Clustered regularly interspaced palindromic repeats” (CRISPR/cas) loci, for example, recognize invading DNA and confer bacterial adaptive immunity to phage infection[3]. Other strategies to avoid HGT include restriction modification (R-M) systems, which most likely evolved in order to avoid uptake of foreign DNA from sources other than the same or related bacterial species[1,4–6]. However, in many pathogenic bacteria including the major human pathogen *Staphylococcus aureus*, particular phage-mediated transduction is probably the most efficient and important mechanism to exchange genetic information[7,8]. Typically, *S. aureus* benefits from phage-mediated HGT events, since many staphylococcal phages mobilize resistance plasmids, genomic islands or other genomic loci with determinants of bacterial virulence[9,10], thus substantially contributing to the evolution, pathogenicity, and global spread of this pathogen. Hence, protective mechanisms, which interfere with or even completely prevent phage infection and phage-mediated HGT events, can appear disadvantageous and maintain pathogens such as *S. aureus* in an evolutionary “dead-end”. Such a scenario is probably a reason for the emergence of phylogenetically isolated branches, as reported recently for the unique *S. aureus* lineage sequence type (ST) 395, which completely changed the phage adsorption receptor properties rendering it resistant from HGT with other *S. aureus* lineages[11,12]. However, such dramatic changes in the phage receptor properties are probably very rare among *S. aureus* clones and do not

represent a frequent strategy to prevent phage adsorption or other phage-mediated HGT events.

Apart from ST395 isolates, which synthesize a unique glycerol-phosphate (GroP) WTA substituted with D-alanine and α-O-N-Acetylgalactosamine (GalNAc)[11,12], most *S. aureus* clones synthesize a ribitol-phosphate (RboP) WTA repeating unit substituted with three tailoring modifications, D-alanine, α-O-N-acetylglucosamine (GlcNAc), and β-O-GlcNAc[13,14]. The GlcNAc moieties are attached to RboP by two independent enzymes, the α-O-GlcNAc WTA glycosyltransferase TarM[15], and the β-O-GlcNAc WTA transferase TarS[16]. Most S. aureus phages and phage-related *S. aureus* pathogenicity island (SaPI) particles target these WTA O-GlcNAc moieties for adsorption and subsequent infection[11,15–17]. Apparently, the stereochemical linkage of WTA glycosylation is dispensable for the phage infection process since strains lacking one of the two WTA glycosyltransferases are still phage or SaPI-particle susceptible[11,16]. In contrast, staphylococcal *Myoviridae* simply require WTA polymers, regardless of the polyol type or WTA O-GlcNAcylation[11,12,17]. Nevertheless, since WTA polymers have many other crucial functions in *S. aureus* pathogenesis and resistance[13,14], most staphylococcal phages seem to be well-adapted to a rather conserved and important cell surface molecule, which *S. aureus* presumably does not mutate frequently. Accordingly, phage infection-preventing mutations in WTA biosynthesis genes have not been described so far. Thus, phage-mediated HGT events among *S. aureus* clones frequently occur and are rather beneficial for *S. aureus* evolution and adaptation to changing selection pressures, which is, conversely, also supported by the notion that many *S. aureus* clones if not all (as suggested by a recent in silico study[18]) have lost CRISPR/cas loci, which otherwise disable or even completely block HGT. Accordingly, staphylococcal phage protection mechanisms most likely evolved to prevent phage lysis, caused by lytic but not by transducing or beneficial phages.

Here, a novel strategy of *S. aureus* is described to prevent adsorption and infection of Podoviridae, a specific class of staphylococcal lytic phages with very short, non-contractile tails. This strain-specific barrier, which was lost by various *S. aureus* lineages during evolution, can completely block the *Podoviridae* infection process thereby providing new insights into bacterial strategies to counteract phage infections.

Results

Infection of *S. aureus* by *Podoviridae* is strain-dependent. Lytic *S. aureus* phages, for example staphylococcal *Myoviridae*, usually have a broad host-range and can even infect other staphylococcal species[11,19]. Accordingly, the broad host-range phages ΦK and Φ812 (*Myoviridae*) infected and lysed nearly all *S. aureus* test strains including strains of dominant MRSA linages, albeit with different potencies (Table 1). However, a collection of another family of lytic staphylococcal phages (*Podoviridae*; here phages Φ44AHJD, Φ66 and ΦP68) failed to infect certain myovirus-susceptible strains, for instance the two American pandemic CA-MRSA clones USA300 (NRS384) and USA400 (MW2), and the HA-MRSA isolate 605, a member of the predominant Asian ST239 lineage (Table 1). Even though some test strains were susceptible to Podoviridae, these phages seem to have a narrower host-range than other lytic staphylococcal phages.

Podovirus-susceptible *S. aureus* strains were found among several clonal lineages suggesting that Podoviridae probably do not require an ST-specific receptor for adsorption and infection, as reported recently for the *S. aureus* ST395-specific phage Φ 187[11,12] (Table 1). In line with this notion, the strains PS44A, PS66, and P68 recommended for propagation of different podoviruses[20] were found to belong to different, unrelated STs, when they were multi locus sequence-typed (MLST) (Table 1).

Thus, staphylococcal *Podoviridae* have a specific host-range different from that of other lytic staphylococcal phages such as *Myoviridae*.

S. aureus strain	Sequence type	*tarM*	*tarS*	Phage susceptibility[b]				
				Myoviridae		*Podoviridae*		
				ΦK	Φ812	Φ44AHJD	Φ66	ΦP68
MW2	1	+	+	+	+	—	—	—
Mu50	5	—	+	(+)	+	+	+	+
USA300	8	+	+	+	+	—	—	—
NRS184	22	—	+	(+)	+	+	+	+
P68	25	—	+	(+)	(+)	+	+	+
UAMS-1	30	+	+	+	+	—	—	—
PS66	39	+	+	+	+	+	+	+
USA600	45	—	+	(+)	—	—	—	—
JH1	105	—	+	+	+	+	+	+
ED133	133	—	—	+	(+)	—	—	—
RF122	151	+	+	+	+	+	+	+
605	239	+	+	(+)	(+)	—	—	—
Col	250	+	+	+	+	—	—	—
PS187[a]	395	—	—	+	+	—	—	—
82086	398	—	+	+	+	+	+	+
PS44A	707	—	+	+	+	+	+	+

Table 1. Lack of *tarM* in *S. aureus* correlates with susceptibility to *Podoviridae*. [a]PS187 synthesizes a poly-glycerol phosphate WTA type modified with α-O-N-Acetylgalactosamine (mediated by the ST395- specific WTA glycosyltransferase TagN[12]). [b]Phage susceptibility was analyzed via soft agar overlay method. Phage susceptibility (+) or resistance is indicated (—). Diminished plaque formation (Φ K, Φ 812) observed for strains Mu50, NRS184, P68, USA600, ED133, and 605 is indicated with a bracketed plus symbol ((+)).

Peptidoglycan-anchored surface proteins are dispensable for host specificity of *Podoviridae*. The specific host-range of *Podoviridae* suggests that these phages might fail to infect and lyse certain *S. aureus* strains due to unique barriers preventing

adsorption, infection, or reproduction. Since the commonly used laboratory and podovirus-resistant *S. aureus* strain RN4220 (see Fig. 1 and Supplementary Fig. S1) lacks R-M systems, prophages, and CRISPR/cas loci previously shown to impede HGT, an intracellular barrier facilitating resistance to *Podoviridae* seems implausible. More likely, alterations in peptidoglycan modifications, for example specific cell-surface exposed molecules such as peptidoglycan-anchored 'microbial surface components recognizing adhesive matrix molecules' (MSCRAMMs), might block adsorption and infection in certain *S. aureus*. However, *S. aureus* RN4220 mutants and mutants derived from the clinical CA-MRSA isolate USA300 lacking functional surface proteins (Δ *srtA*) were resistant to *Podoviridae* indicating that factors other than MSCRAMMs interfere with the podovirus infection process (Supplementary Fig. S1). Thus, *S. aureus* peptidoglycan-anchored surface proteins do not influence the unusual host-range of staphylococcal *Podoviridae*.

The *S. aureus* α-O-GlcNAc WTA glycosyltransferase TarM prevents the lytic activity of *Podoviridae*. Because all studied staphylococcal phages require WTA polymers or O-GlcNAcylated WTA polymers for adsorption and infection[17], adsorption of *Podoviridae* to their designated cell surface receptors may also be influenced by WTA polymers. Of note, all podovirus-susceptible strains were simultaneously susceptible to the WTA-dependent phages Φ K and Φ 812, which excludes that podovirus-susceptible strains fail to produce WTA polymers (Table 1). In line with this assumption, *Podoviridae* still failed to adsorb to and infect *S. aureus* RN4220 or USA300 mutants lacking either WTA (Δ *tagO*) or WTA glycosylation (Δ *tarM* Δ *tarS*) (Fig. 1a,b).

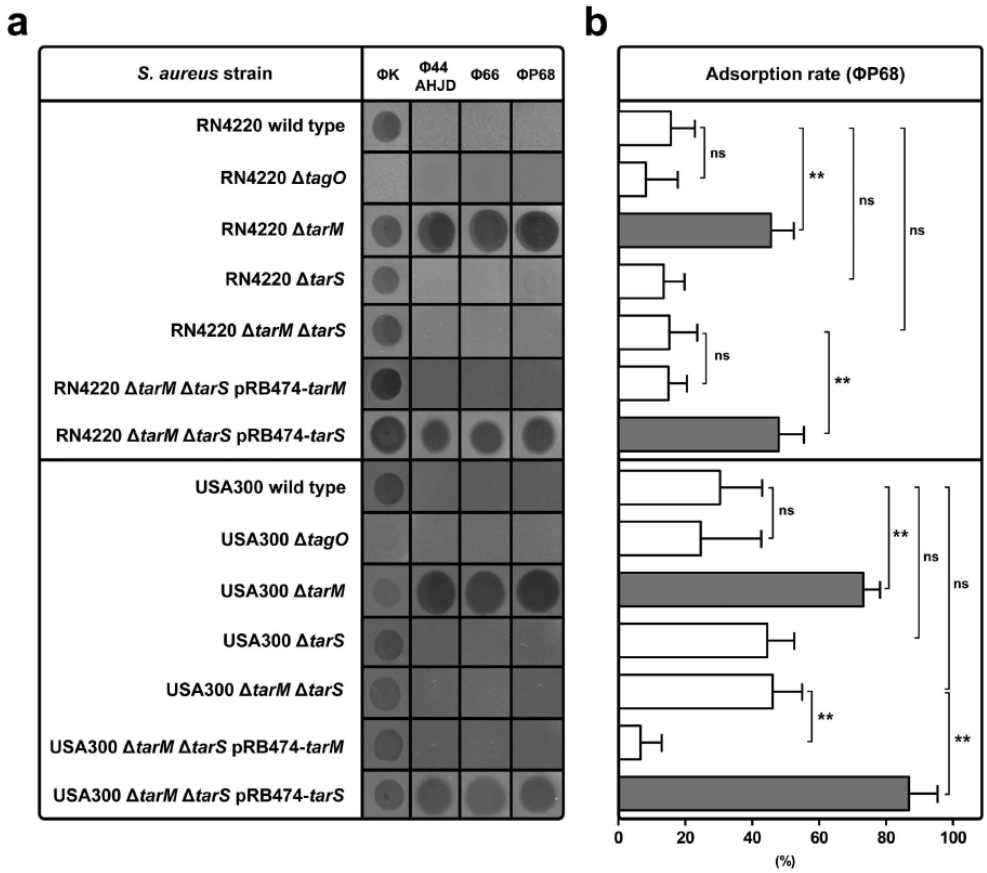


Figure 1. The α-O-GlcNAc WTA glycosyltransferase TarM protects *S. aureus* from the lytic activity of *Podoviridae*. (a) *S. aureus* RN4220 and USA300 susceptibility to the broad-host-range lytic phage ΦK (*Myoviridae*), and to the lytic phages Φ44AHJD, Φ66 and ΦP68 (*Podoviridae*) was analyzed using a softagar overlay approach. A representative experiment is shown. (b) Podovirus ΦP68 adsorption rates (%) to *S. aureus* RN4220 and USA300 variants. *S. aureus* wild type and strains lacking WTA (Δ tagO), WTA α -O-GlcNAcylation (Δ *tarM*), WTA β-O-GlcNAcylation (Δ *tarS*), WTA glycosylation (Δ *tarM* Δ *tarS*), and the complemented mutants (Δ *tarM* Δ *tarS* pRB474-*tarM*, Δ *tarM* Δ *tarS* pRB474-*tarS*) are indicated. Values are given as means and standard deviations (SD, n = 3). Statistical significant differences calculated by oneway ANOVA with Bonferroni's multiple comparison test are indicated: not significant (ns), $P > 0.05$; *$P < 0.05$, **$P < 0.01$.

While well-studied WTA-GlcNAc dependent *S. aureus* phages such as phage Φ11 do not seem to require a specific stereochemistry of WTA O-GlcNAc for infection[16] the tested podoviruses exhibited an unexpected preference for TarS-glycosylated but not TarM-glycosylated WTA. Strikingly, lack of WTA α-O-GlcNAcylation (Δ*tarM*) resulted in dramatically increased binding capacities of phage ΦP68 and rendered strain RN4220Δ*tarM* highly susceptible to podovirus infection (Fig 1a,b). In contrast, lack of *tarS* did not lead to phage susceptibility of RN4220 (Fig. 1a). Complementation of the WTA-glycosylation deficient Δ*tarM*Δ*tarS* mutant with one of the two *S. aureus* WTA glycosyltransferases TarM or TarS demonstrated that, (i) *Podoviridae* require TarS-mediated WTA β-O-GlcNAcylation, but (ii) are inhibited by TarM-mediated WTA β-O-GlcNAcylation (Fig 1a,b). Similar results were obtained for *S. aureus* USA300 strongly suggesting that TarM diminishes the adsorption and infection of *Podoviridae* to *S. aureus* (Fig. 1a,b). Because TarM is an intracellular protein it appears highly unlikely that it interferes with podovirus binding directly but impedes podovirus binding by α-O-GlcNAcylated WTA.

Thus, the α -O-GlcNAc WTA glycosyltransferase TarM prevents the adsorption and infection by staphylococcal *Podoviridae.*

Lack of *tarM* correlates with susceptibility to *Podoviridae*. In order to confirm the inhibitory effect of TarM on podovirus susceptibility, genomes of *S. aureus* test strains were screened for the presence or absence of the genes encoding WTA glycosyltransferases TarM and TarS via PCR or BLASTN of available genomes21. Most strains contained *tarS* except for strains PS187, which produce an entirely different type of WTA[11,12], and ED133, which does not encode any of the so far described WTA glycosyltransferases (Table 1). In contrast, several strains lacked *tarM*. As proposed, most *tarM*- plus *tarS*-encoding *S. aureus* strains were podovirus-resistant (Table 1). Conversely, *S. aureus* strains exclusively encoding *tarS*
and even other staphylococcal species such as *Staphylococcus xylosus* or *Staphylococcus equorum*, which encode *tarS* homologues with high similarity, but

lack *tarM*, were susceptible indicating that *Podoviridae* specifically sense β-O-GlcNAcylated WTA (Table 1 and Supplementary Fig. S2). In line with this, the designated podovirus propagation strains PS44A (Φ44AHJD) and P68 (ΦP68) exclusively encoded *tarS* (Table 1). However, strain PS66 (Φ66) encoded both WTA glycosyltranserases, TarM and TarS, which did not align with the assumption that *tarM* interferes with podovirus susceptibility. Nevertheless, even though *tarM* was expressed at good levels during logarithmic growth phase, *tarS* was significantly higher expressed than *tarM* during early growth stages, which probably promotes the infection by *Podoviridae* (Supplementary Fig. S3). Moreover, the *S. aureus* PS66 *tarM* gene was sequenced and found to contain two non-synonymous point mutations (Q453K and A464E), which may compromise the TarM function and capacity to interfere with podovirus infection (Fig. 2a). Indeed, podovirus resistance of RN4220Δ*tarM*, whose WTA contains only β-O-GlcNAc could be restored completely by complementation with a wild-type *tarM* but only partially by the mutated *tarM* (Fig. 2b). In addition, deletion of *tarS* in PS66 resulted in drastically reduced binding capacity of ΦP68 and rendered PS66 resistant to *Podoviridae* (Supplementary Fig. S4) suggesting that podovirus sensitivity of PS66 is linked to *tarS*-mediated β-O-GlcNAcylated WTA and to a strain-specific dysfunction of TarM.

Next, *tarM* was expressed in various podovirus-susceptible strains, including the Φ44AHJD and Φ66 propagation strains PS44A and PS66. Even at very high phage titers, expression of *tarM* rendered most susceptible strains completely resistant, confirming the importance of *tarM* in diminishing infection by staphylococcal *Podoviridae* (Fig. 3). In addition, the expression of a plasmid-born copy of *tarM* in strain PS66 also caused complete resistance to *Podoviridae*, further suggesting that the *tarM* gene of PS66 is most likely non-functional or less active (Fig. 3).

Thus, *Podoviridae* require β-O-GlcNAcylated WTA but cannot infect *S. aureus* with α-O-GlcNAcylated WTA.

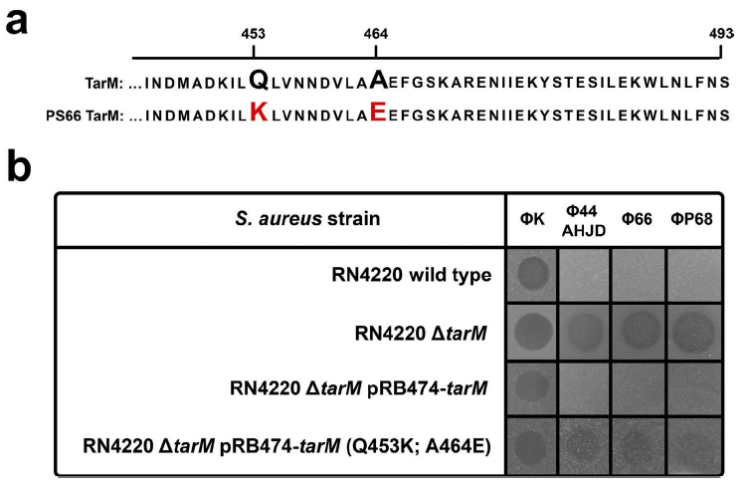


Figure 2. Point mutations in TarM render Φ66 propagation strain PS66 susceptible to Podoviridae. (a) A sequence alignment of wild-type TarM and PS66 TarM is shown. Position of mutations (Gln-453 with Lys; Ala-464 with Glu) and the end of the open reading frame (493) are indicated. (b) *S. aureus* RN4220 susceptibility to the broad host-range lytic phage ΦK (*Myoviridae*), and to the lytic phages Φ44AHJD, Φ66, and ΦP68 (*Podoviridae*) was analyzed using a soft-agar overlay approach. *S. aureus* RN4220 wild type and strains lacking WTA α-O-GlcNAcylation (Δ*tarM*), and the complemented mutants (Δ*tarM* pRB474-*tarM*, Δ*tarM* pRB474-*tarM* (Q453K; A464E) are indicated. A representative experiment is shown.

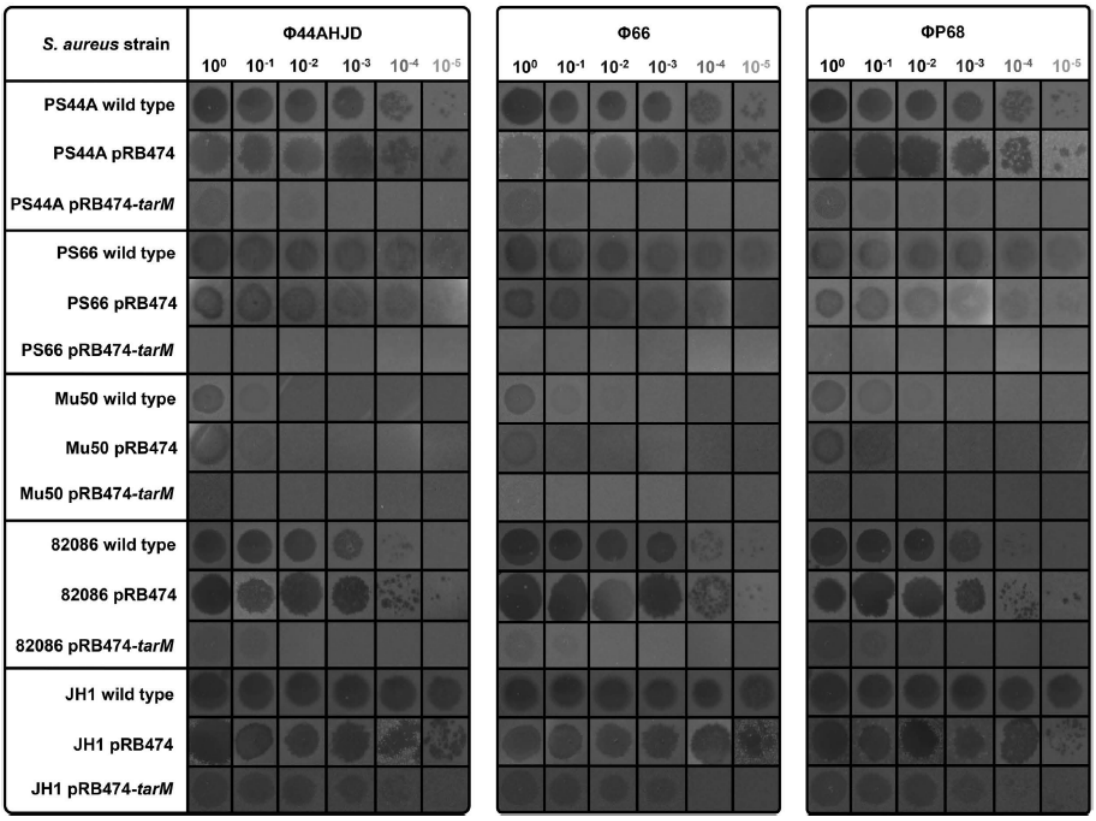


Figure 3. Ectopic expression of TarM protects podovirus-susceptible *S. aureus* against *Podoviridae*. The α-O-GlcNAc WTA glycosyltransferase TarM was ectopically expressed in various *tarM*-lacking and podovirus-susceptible *S. aureus* strains, and the phage susceptibility using a phage panel encompassing the lytic phages Φ44AHJD, Φ66 and ΦP68 (*Podoviridae*) was analyzed using a soft-agar overlay approach. Various dilutions of phage lysates, *S. aureus* wild type strains (*tarS* positive, but *tarM* negative (or encoding a mutated *tarM*, strain PS66)), and engineered strains expressing *tarM* (pRB474-*tarM*), or empty plasmid control (pRB474) are indicated. A representative experiment is shown.

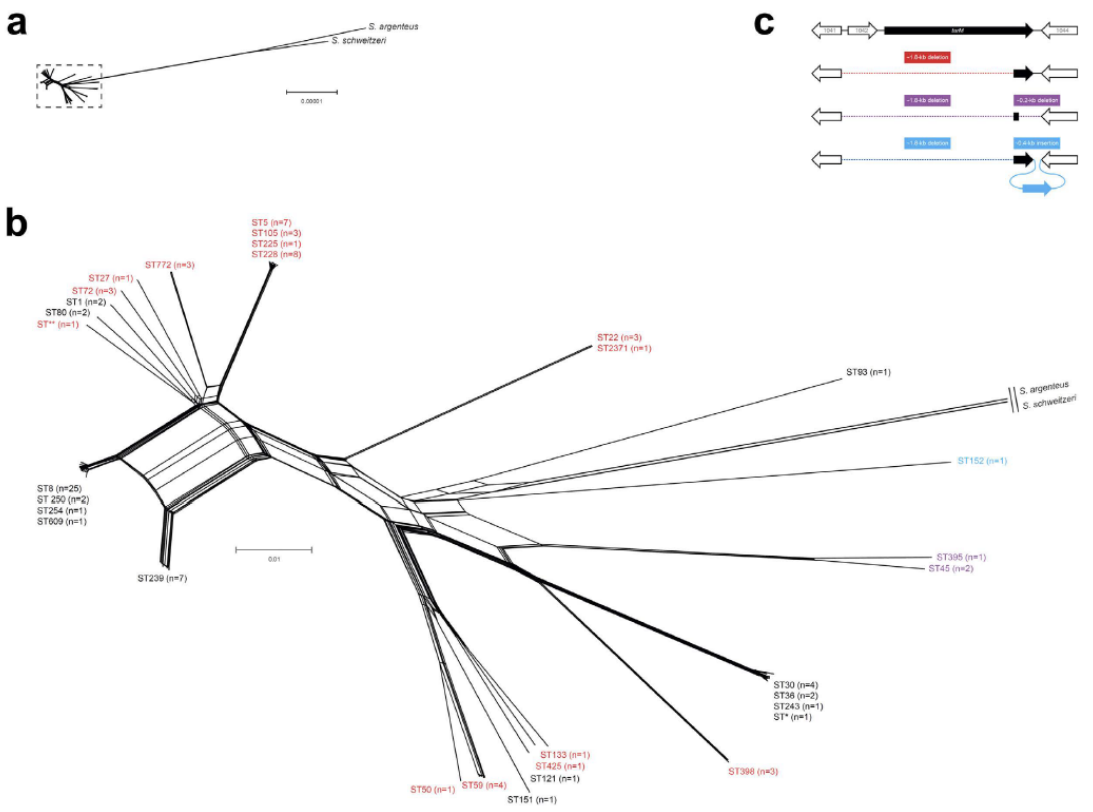


Figure 4. Phylogenetic distribution of *tarM* reveals an ancient origin in other *staphylococci* and vertical inheritance during *S. aureus* evolution. (a,b) Phylogenetic network representing the inferred relationship of 98 *S. aureus* strains and two closely related species, *S. argenteus* and *S. schweitzeri*. Strains are indicated by their multilocus sequence types (STs). ST* and ST** are single-locus variants of ST30 and ST1148, respectively. Strains encoding *tarM* are indicated in black, while strains lacking *tarM* are indicated in red, purple, and blue. (c) Genetic organization of the *tarM* region in *S. aureus*. The intact *tarM* region is shown in the upper cluster. Gene locus numbers refer to *S. aureus* strain COL (GenBank accession no. CP000046). Lower clusters indicate distinct deletion events involving *tarM*.

Tracking the evolution of TarM reveals an ancient origin in other staphylococcal species and vertical inheritance during *S. aureus* evolution. TarM is encoded outside of the *S. aureus* WTA gene clusters but does not appear to be encoded on a mobile genetic element[22]. Nevertheless, it is tempting to assume that it has been acquired by *S. aureus* at some point in evolution to interfere with podovirus infection.

To track the emergence of TarM in *S. aureus*, the genome sequences of 98 *S. aureus* strains including those of most *S. aureus* laboratory test strains used in this study were obtained to infer their genetic relatedness (Fig. 4a,b). Of note, the presence of *tarM* in the most deeply branching *S. aureus* isolates MSHR1132 and FSA084, which were recently proposed as novel staphylococcal species *Staphylococcus argenteus* sp. nov. and *Staphylococcus schweitzeri* sp. nov.[23], revealed that the presence of *tarM* is probably an ancient genetic trait of *S. aureus* (Fig. 4a). Still, homologues of *tarM* are also encoded by certain coagulase-negative *staphylococci* (e.g. specific *S. epidermidis* isolates) and even by non-staphylococcal species such as *Exiguobacterium oxidotolerans* and *Tetragenococcus halophilus*. Thus, the early evolution of *tarM* probably involved an ancient HGT event to the last common ancestor of contemporary *S. aureus* clones, further supported by the notion that *tarM* is flanked by a gene possibly related to conjugation (SACOL1042) (Fig. 4c). However, at a later stage of *S. aureus* evolution, different types of genetic rearrangements occurred in emerging phylogenetic branches such as CC5 or CC398, leading to a deletion of *tarM*, which rendered these podovirus-susceptible (Fig. 4c).

Discussion

Staphylococcal *Podoviridae* infect an unusually wide panel of staphylococcal species but remain avirulent for certain *S. aureus* lineages probably as a result of the activity of the α -O-GlcNAc WTA glycosyltransferase TarM. In *tarM*-encoding strains, WTA polymers are probably glycosylated preferentially with α-O-GlcNAc, suggesting that TarM might be more active than TarS. Consequently, TarS-mediated β-O-GlcNAcylation is probably affected by the activity of TarM, thus preventing the adsorption and infection of *Podoviridae*. Even though it cannot be excluded that TarM potentially has additional and undiscovered functions, which may interfere with the adsorption or infection process, the drastically increased adsorption of ΦP68 in isogenic Δ *tarM* mutants suggests that α-O-GlcNAcylated WTA prevents the adsorption of *Podoviridae* to *S. aureus*. Nevertheless, one of the designated podovirus propagation strains (PS66) encoded both WTA glycosyltransferases suggesting that certain strains, despite encoding *tarM*, are potentially podovirus-susceptible. Here, TarM might be non-functional, dis-regulated, or mutated as observed in PS66, and cannot interfere with the activity of TarS. Nevertheless, this TarM-mediated phenomenon limits the host-range of *Podoviridae*, and thus, their therapeutic potential compared to other lytic staphylococcal phages such as *Myoviridae*.

Apart from this, it remains intriguing as to why certain strains and lineages have lost *tarM* during evolution to become podovirus-susceptible. Since both *S. aureus* and *S. aureus*-like species such as *S. schweitzeri* and *S. argenteus* encode *tarM* and *tarS*, and many human-associated *S. aureus* lineages have lost *tarM* during evolution, it can be assumed that *tarM* is probably not essential for continued adaptation to the human host. This is in agreement with the observation that both types of WTA O-GlcNAcylation, can mediate *S. aureus* binding to nasal epithelial cells and thus nasal colonization[24]. Also, human sera contain preferentially serum antibodies directed against TarS-dependent β -O-GlcNAcylated WTA, but not against TarM-mediated α -O-GlcNAcylated WTA[25], suggesting that *tarM* may be down-regulated or less immunogenic than β -O-GlcNAcylated WTA during infections. It can be assumed

that some *S. aureus* lineages did not eliminate *tarM* because WTA α-O-GlcNAcylation may provide *S. aureus* with a fitness benefit, whose basis remains to be identified in the future.

However, bearing *tarM* and TarM-mediated α-O-GlcNAcylated WTA protects *S. aureus* at least against the lytic activity of staphylococcal *Podoviridae* via a modification of the designated phage adsorption receptor. Such alterations of cell-surface structures serving as viral receptors are only one of many bacterial strategies to counteract phage infection and have also been described for other bacterial species[26–28], but does not seem a general strategy of *S. aureus* to avoid phage adsorption and infection. Since other lytic staphylococcal phages such as *Myoviridae* are capable of infecting *tarM*-encoding *S. aureus* isolates, prevention of podovirus infection could be the result of a highly specific WTA-dependent mechanism in *S. aureus*, presumably as the result of adaptation to specific podovirus-rich environmental niches. In addition, altered phage-receptor binding proteins may easily change the host-range of *Podoviridae* to render *tarM*-bearing clones susceptible. Whereas bacterial phage resistance mechanisms such as CRISPR interference appear more efficient and widespread in prokaryotes these can also be bypassed, for example, by CRISPR-evading phages[29] suggesting that host-virus interaction is a constantly evolving process.

Methods

Bacterial strains and growth conditions. All bacterial strains used in this study are listed in Supplementary Table S1. Unless otherwise noted, bacteria were grown in basic medium (BM) (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 0.1% K2HPO4, 0.1% glucose) or lysogeny broth (Becton Dickinson) supplemented with appropriate antibiotics (Chlorampenicol 10 μ g/ml, Ampicillin 100 μ g/ml).

Molecular genetic methods. *S. aureus* RN4220 and USA300 Δ*tarM*, Δ*tarS*, Δ*tarM* Δ*tarS*, and Δ*tagO* deletion mutants were described elsewhere[11,16,24]. For the construction of marker-less RN4220 Δ*srtA* mutant, or a PS66 Δ*tarS* mutant, the previously described *E. coli*/*S. aureus* shuttle vectors pIMAY or pKOR1 were used[30,31]. The corresponding primers are listed in Supplementary Table S2. Gene disruption by using pKOR1 or pIMAY was performed as described before[30,31]. Briefly, pKOR1-*tarS*, or pIMAY-*srtA* were isolated from an appropriate *E. coli* strain, and transformed into electrocompetent *S. aureus* RN4220 cells (and reisolated and transformed into PS66). Electroporation conditions were described before[11]. Knock-out plasmids were integrated onto the *S. aureus* genome at the permissive temperatures (37 °C, pIMAY; 43 °C, pKOR1) and in the presence of chloramphenicol (10 μg/ml). Counterselection was performed by using anhydrotetracycline (1 μg/ml). Resulting colonies were patched onto BM agar plates with and without chloramphenicol (10 μg/ml) and screened for plasmid loss. Gene deletion was confirmed via PCR in chloramphenicol-sensitive colonies. For complementation studies (or *tarM* expression in *tarM*-lacking strain backgrounds), the previously described *E. coli*/*S. aureus* shuttle vector pRB474 was used[32]. pRB474-*tarM* (Q453K; A464E) has been described elsewhere (formerly pRB474-H-*tarM*)[15].

PCR-typing, sequencing, and multiple locus sequence typing (MLST). For verification (and sequencing) of *tarM* and *tarS* in *S. aureus* genomes, PCR analysis using primers listed in Supplementary Table S2 was used. MLST typing of podovirus propagation strains PS44A, PS66 and P68 was performed as described previously using published primers[33].

Experiments with phages. All phages used in this study are listed in Supplementary Table S1. Phages were propagated on *S. aureus* strains P68 or RN4220Δ*tarM* (Φ44AHJD, Φ66 and ΦP68), or RN4220 wild type (ΦK, Φ812) as described previously[34]. Briefly, the cognate *S. aureus* host strains were grown overnight at 37 °C in BM and diluted in phage-containing lysates (approximately 1×10^9 plaque forming units (PfU) per milliliter; titrated on cognate host strains) to a final optical density OD 600 nm of 0.4. Subsequently, CaCl2 was added to a final concentration of 4 mM. The bacteria/phage mixture was incubated for 30 min at 37 °C without agitation and afterwards for at least 3 h at 30 °C with mild agitation until complete lysis occurred. In order to remove cell debris, the lysate was centrifuged for 10 min (5,000 g, 4 °C). Lysates were filter-sterilized (0.22 μ m) and stored at 4 °C. Phage susceptibility was analyzed as described elsewhere[17]. Briefly, a phage panel encompassing the broad host-range phages ΦK and Φ812 (*Myoviridae*), and three serogroup G phages Φ44AHJD, Φ66 and ΦP68 (*Podoviridae*) were used. 6 μ l (approximately 1×10^9 PfU/ml, or appropriate dilutions) of freshly propagated phage lysates were spotted onto bacterial lawns using the soft-agar overlay method described by Xia *et al.*[17]. Phage adsorption to *S. aureus* strains was analyzed as described previously[17]. Briefly, the phage adsorption rate was analyzed using a multiplicity of infection (MOI) of 0.01 for phage ΦP68. Adsorption rate (%) was calculated by determining the number of unbound PfU in the supernatant and subtracting from the total number of input PfU as a ratio to the total number of input PfU.

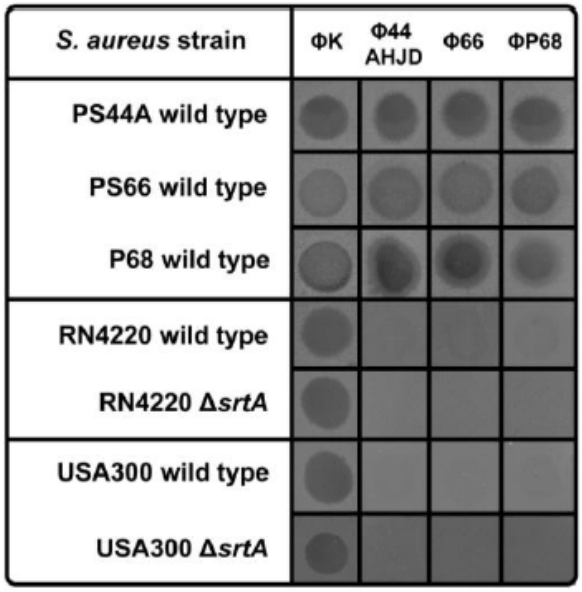
Phylogenetic analysis. The chromosomes of all *S. aureus* and *S. argenteus* and *S. schweitzeri* labelled as complete were obtained from GenBank (Supplementary Table S3) and aligned against the chromosome of *S. aureus* CC45 strain CA-347 (GenBank accession ID NC_021554) after identification and deletion of duplicated regions using MUMmer v 3.22[35]. The 98 publicly available genomes were aligned using MUMmer. Based on the identified core of ~1,9 Mb (67%) among all strains, a total of 312,427 SPNs was identified, from which the phylogenetic relationship was inferred using the NeighbourNets algorithm in SplitsTree v4.13.1[36].

RNA isolation and preparation. RNA was isolated as described previously[24]. Briefly, BM over-night cultures were diluted in BM. Bacteria were grown at 37 °C until lag, log, or stationary growth phases. Subsequently, bacteria were harvested and resolved in 1 ml TRIzol (Invitrogen/Life Technologies, Karlsruhe, Germany). Next, bacteria were mechanically disrupted by using a FastPrep24 homogenizer (MP Biomedicals) (2 cycles, 20 sec. at 6500 rpm each, with 0.5 ml Zirconia/Silica beads (0.1 mm in diameter; Carl-Roth, Karlsruhe, Germany)). Samples were either stored at − 80 °C or subsequently used for further preparation. To each sample, 200 µl chloroform was added and samples were thoroughly mixed for 60 s, and incubated for 3 min at room temperature. Samples were centrifuged at 4 °C (12,000 . g, 15 min) and the supernatant was mixed with 500 µl isopropanol. Next, samples were incubated for 10 min at room temperature and centrifuged (12,000 . g, 30 min, 4 °C). Each pellet was washed with 500 µ l ethanol (70%) and the sample was centrifuged (7,500 . g, 5 min, 4 °C). Finally, the pellet was air-dried and dissolved in 50 µl nuclease-free water. After incubation at 55 °C for 10 min, the sample was mixed well for 4 min. 5 µ g RNA was digested with DNAse I (Roche) and stored at − 80 °C.

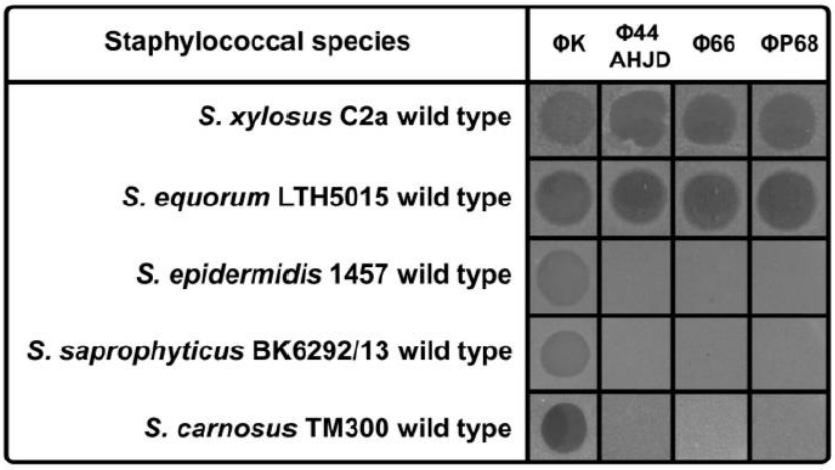
Quantitative real time PCR (qRT-PCR). qRT-PCR was performed as described previously[24]. Briefly, RNA was transcribed into cDNA and qRT-PCR was performed according to the manufactures instructions using the Brilliant II SYBR© Green 1-Step Master Mix (Agilent). Relative quantifications were analyzed by using Roche's LightCylcer480II. Transcription levels of target genes analyzed in this study were normalized against the expression of the housekeeping gene *gyrB.* All primers used for qRT-PCR are listed in Supplementary Table S2.

Statistical analysis. Statistical analysis was performed using GraphPad Prism (GraphPad Software, Inc., La Jolla USA, Version 5.04). Statistically significant differences were calculated by using appropriate statistical methods as indicated. P values < 0.05 were considered significant.

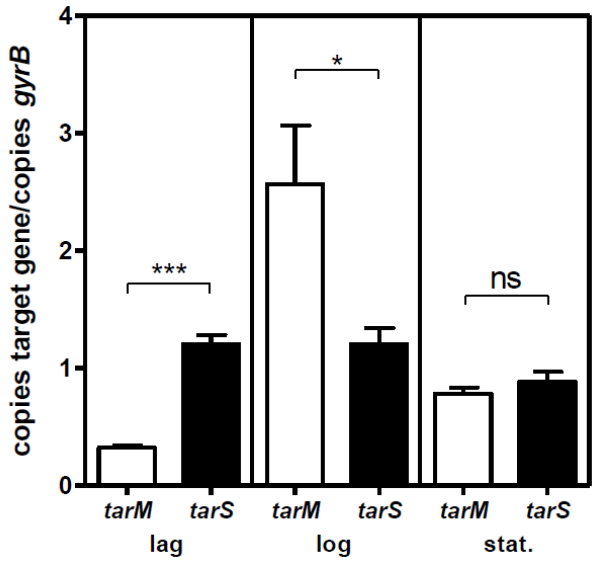
Supplementary Material



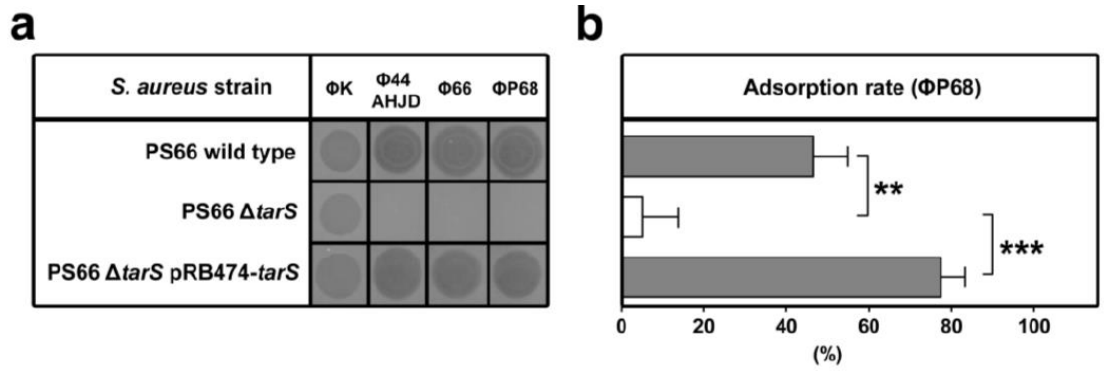
Supplementary Figure S1 – Impact of peptidoglycan-anchored surface proteins on host-specificity of *Podoviridae*. *S. aureus* RN4220 and USA300 susceptibility to the broad host- range lytic phage ΦK (*Myoviridae*), and to the lytic phages Φ44AHJD, Φ66 and ΦP68 *Podoviridae*) was analyzed using a soft-agar overlay approach. *S. aureus* podovirus propagation strains (PS44A, PS66, and P68), *S. aureus* wild type, and mutants lacking peptidoglycan-anchored surface proteins (Δ*srtA*) are indicated. A representative experiment is shown.



Supplementary Figure S2 – Susceptibility of selected staphylococcal species to *Podoviridae*. Susceptibility of selected staphylococcal species to the broad host-range lytic phage ΦK (*Myoviridae*), and to the lytic phages Φ44AHJD, Φ66 and ΦP68 (*Podoviridae*) was analyzed using a soft-agar overlay approach. A representative experiment is shown.



Supplementary Figure S3 – qRT-PCR analysis of WTA glycosyltransferases in Φ66 propagation strain PS66. mRNA was isolated from lag-, log-, or stationary (stat.)-phase- grown bacteria. Values are given as means and standard deviations (SD; n = 3). Statistically significant differences calculated using an unpaired two-tailed Student's t test are indicated as follows: ns (not significant), $P > 0.05$; *, $P < 0.05$; ***, $P < 0.001$.



Supplementary Figure S4 – Lack of the β-O-GlcNAc WTA glycosyltransferase TarS renders Φ66 propagation strain PS66 resistant to *Podoviridae*. (a) *S. aureus* PS66 susceptibility to the broad host-range lytic phage ΦK (*Myoviridae*), and to the lytic phages Φ44AHJD, Φ66 and ΦP68 (*Podoviridae*) was analyzed using a soft-agar overlay approach. A representative experiment is shown. **(b)** Podovirus ΦP68 adsorption rate (%) to *S. aureus* PS66 variants. *S. aureus* wild type, the strain lacking WTA β-O-GlcNAcylation (Δ*tarS*), and the complemented Δ*tarS* mutant (Δ*tarS* pRB474-*tarS*) are indicated. Values are given as means and standard deviations (SD, n = 3). Statistical significant differences calculated by one-way ANOVA with Bonferroni's multiple comparison test are indicated: not significant (ns), $P > 0.05$; *, $P < 0.05$, **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

Supplementary Table S1 - Bacterial strains and phages used in this study.

Bacterial strain or phage	Description	Source
E. coli TOP10	One Shot® TOP10 chemically competent *E. coli*	Invitrogen
E. coli DB 3.1 pKOR1	DB3.1 strain, bears pKOR1 plasmid	[30]
E. coli DC10B pIMAY	DH10B Δ*dcm*; Dam methylation only, bears pIMAY plasmid	[31]
S. aureus RN4220	Wild type, deficient in restriction, capsule, and prophage	[37]
S. aureus RN4220 Δ*tagO*	RN4220 Δ*tagO*	[17]
S. aureus RN4220 Δ*tarM*	RN4220 Δ*tarM*	[11]
S. aureus RN4220 Δ*tarS*	RN4220 Δ*tarS*	[11]
S. aureus RN4220 Δ*tarM* Δ*tarS*	RN4220 Δ*tarM* Δ*tarS*	[11]
S. aureus RN4220 Δ*tarM* Δ*tarS* pRB474-*tarM*	RN4220 Δ*tarM* Δ*tarS* complemented with *tarM*	[11]
S. aureus RN4220 Δ*tarM* Δ*tarS* pRB474-*tarS*	RN4220 Δ*tarM* Δ*tarS* complemented with *tarS*	[11]
S. aureus RN4220 Δ*srtA*	RN4220 Δ*srtA*	This study
S. aureus USA300	Wild type, NRS384, ST8, CA-MRSA	NARSA strain collection
S. aureus USA300 Δ*tagO*	USA300 Δ*tagO*	[24]
S. aureus USA300 Δ*tarM*	USA300 Δ*tarM*	[24]
S. aureus USA300 Δ*tarS*	USA300 Δ*tarS*	[16]
S. aureus USA300 Δ*tarM* Δ*tarS*	USA300 Δ*tarM* Δ*tarS*	[24]
S. aureus USA300 Δ*tarM* Δ*tarS* pRB474-*tarM*	USA300 Δ*tarM* Δ*tarS* complemented with *tarM*	[24]
S. aureus USA300 Δ*tarM* Δ*tarS* pRB474-*tarS*	USA300 Δ*tarM* Δ*tarS* complemented with *tarS*	[24]
S. aureus USA300 Δ*srtA*	USA300 Δ*srtA*	[24]
S. aureus PS44A	Wild type, PS44A, NCTC 8369, ST707, designated propagation strain for Φ44AHJD	NCTC collection
S. aureus PS66	Wild type, PS66, NCTC 8288, ST39, designated propagation strain for Φ66	Obtained from Udo Bläsi, Vienna
S. aureus PS66 Δ*tarS*	PS66 Δ*tarS*	This study
S. aureus PS66 Δ*tarS* pRB474-*tarS*	PS66 Δ*tarS* complemented with *tarS*	This study
S. aureus P68	Wild type, P68, ST25, designated propagation strain for ΦP68	Obtained from Udo Bläsi, Vienna
S. aureus RF122	Wild type, bovine isolate, ST151	[38]
S. aureus USA600	Wild type, NRS22, ST45	NARSA strain collection
S. aureus ED133	Wild type, ovine isolate, ST133	[39]
S. aureus Col	Wild type, clinical isolate, ST250	[40]
S. aureus PS187	Wild type, PS187, ST395	[41]
S. aureus MW2	Wild type, MW2, ST1	[42]
S. aureus Mu50	Wild type, Mu50, ST5	[43]
S. aureus NRS184	Wild type, NRS184, ST22	NARSA strain collection
S. aureus JH1	Wild type, JH1, ST105	[44]
S. aureus 605	Wild type, 605, ST239	[45]
S. aureus 82086	Wild type, 82086, ST398	[24]
S. aureus 82086 pRB474-*tarM*	Wild type, 82086, bears pRB474-*tarM*	This study

S. aureus 82086 pRB474	Wild type, 82086, bears pRB474	This study
S. aureus Mu50 pRB474-*tarM*	Wild type, Mu50, bears pRB474-*tarM*	This study
S. aureus Mu50 pRB474	Wild type, Mu50, bears pRB474	This study
S. aureus JH1 pRB474-*tarM*	Wild type, JH1, bears pRB474-*tarM*	This study
S. aureus JH1 pRB474	Wild type, JH1, bears pRB474	This study
S. aureus PS44A pRB474-*tarM*	Wild type, PS44A, bears pRB474-*tarM*	This study
S. aureus PS44A pRB474	Wild type, PS44A, bears pRB474	This study
S. aureus PS66 pRB474-*tarM*	Wild type, PS66, bears pRB474-*tarM*	This study
S. aureus PS66 pRB474	Wild type, PS66, bears pRB474	This study
S. xylosus C2a	Wild type, human skin isolate, DSM20267	[46]
S. equorum LTH5015	Wild type	Obtained from Friedrich Götz, Tuebingen
S. epidermidis 1457	Wild type, clinical isolate	[47]
S. saprophyticus BK6292/13	Wild type, clinical isolate	Obtained from Holger Rohde, Hamburg
Phage ΦK	*Myoviridae*, Serogroup D	[48]
Phage Φ812	*Myoviridae*, Serogroup D	[19]
Phage Φ44AHJD	*Podoviridae*, Serogroup G	Obtained from Udo Bläsi, Vienna
Phage Φ66	*Podoviridae*, Serogroup G	Obtained from Udo Bläsi, Vienna
Phage Φ68	*Podoviridae*, Serogroup G	Obtained from Udo Bläsi, Vienna

Supplementary Table S2 – Oligonucleotides used in this study.

Primer	Sequence	Application	Reference
tarM-up	ATGAAAAAAATATTTATGATGGTACATGAGTTAGA	PCR-typing *tarM*	[24]
tarM-dn	TTAGCTATTGAAAAGATTTAACCATTTTTCTAATA	PCR-typing *tarM*	[24]
tarS-up	ATGATGAAATTTTCAGTAATAGTTCCAACATACAA	PCR-typing *tarS*	[24]
tarS-dn	TTATTTTAGCGAGTAAGTCATATGTGCAGT	PCR-typing *tarS*	[24]
tarM-for2	TAATGCTAATAATGGTGCTG	qRT-PCR	[16]
tarM-rev2	GGTCCATCACAAATCATAAT	qRT-PCR	[16]
tarS-for2	CACGAAACAAGAAGCACA	qRT-PCR	[16]
tarS-rev2	TGATTACCAACACGCACT	qRT-PCR	[16]
gyrBF	GGTGGCGACTTTGATCTAGC	qRT-PCR	[49]
gyrBRv	TTATACAACGGTGGCTGTGC	qRT-PCR	[49]

Supplementary Table S3 – Genome sequences used for phylogenetic analysis.

Species	Strain	Accession	Total length	ST	*tarM*	*tarS*
S. aureus	MW2	BA000033	2820462 bp	1	+	+
S. aureus	MSSA476	BX571857	2799802 bp	1	+	+
S. aureus	Mu3	AP009324	2880168 bp	5	-	+
S. aureus	Mu50	BA000017	2878529 bp	5	-	+
S. aureus	N315	BA000018	2814816 bp	5	-	+
S. aureus	ED98	CP001781	2824404 bp	5	-	+
S. aureus	502A	CP007454	2764699 bp	5	-	+
S. aureus	ECT-R 2	FR714927	2729540 bp	5	-	+
S. aureus	M0628	KB821506	2828436 bp	5	-	+
S. aureus	NCTC 8325	CP000253	2821361 bp	8	+	+
S. aureus	USA300_FPR3757	CP000255	2872769 bp	8	+	+
S. aureus	USA300_TCH1516	CP000730	2872915 bp	8	+	+
S. aureus	VC40	CP003033	2692570 bp	8	+	+
S. aureus	USA300-ISMMS1	CP007176	2921008 bp	8	+	+
S. aureus	2395 USA500	CP007499	2955646 bp	8	+	+
S. aureus	M1216	CP007670	2896143 bp	8	+	+
S. aureus	CA15	CP007674	2839253 bp	8	+	+
S. aureus	UA-S391_USA300	CP007690	2872916 bp	8	+	+
S. aureus	29b_MRSA (ATCC BAA-1680)	CP010295	2872768 bp	8	+	+
S. aureus	31b_MRSA (ATCC BAA-1680)	CP010296	2872779 bp	8	+	+
S. aureus	33b (ATCC BAA-1680)	CP010297	2872764 bp	8	+	+
S. aureus	26b_MRSA (ATCC BAA-1680)	CP010298	2872779 bp	8	+	+
S. aureus	25b_MRSA (ATCC BAA-1680)	CP010299	2872781 bp	8	+	+
S. aureus	27b_MRSA (ATCC BAA-1680)	CP010300	2872771 bp	8	+	+
S. aureus	DSM 20231	CP011526	2755072 bp	8	+	+
S. aureus	M1	HF937103	2864125 bp	8	+	+
S. aureus	W48872	KK022849	2859671 bp	8	+	+
S. aureus	T87526	KK027252	2867276 bp	8	+	+
S. aureus	F26088	KK029104	2859631 bp	8	+	+
S. aureus	W15997	KK032093	2860093 bp	8	+	+
S. aureus	M49474	KK072349	2859938 bp	8	+	+
S. aureus	T36111	KK073475	2859914 bp	8	+	+
S. aureus	T83543	KK095312	2860767 bp	8	+	+
S. aureus	NCTC8532	LN831049	2709282 bp	8	+	+
S. aureus	H-EMRSA-15	CP007659	2846320 bp	22	-	+
S. aureus	71A_S11	CP010940	2756431 bp	22	-	+

S. aureus	HO 5096 0412	HE681097	2832299 bp	22	-	+
S. aureus	RKI4	CP011528	2725654 bp	27	-	+
S. aureus	55/2053	CP002388	2756919 bp	30	+	+
S. aureus	FORC_001	CP009554	2886017 bp	30	+	+
S. aureus	UAMS-1	JTJK00000000	2763963 bp	30	+	+
S. aureus	ILRI_Eymole1/1	LN626917	2874302 bp	30	+	+
S. aureus	MRSA252	BX571856	2902619 bp	36	+	+
S. aureus	M0513	KB821413	2932850 bp	36	+	+
S. aureus	CA-347	CP006044	2850503 bp	45	-	+
S. aureus	NRS22	JYAG00000000	2922078 bp	45	-	+
S. aureus	6850	CP006706	2736560 bp	50	-	+
S. aureus	M013	CP003166	2788636 bp	59	-	+
S. aureus	SA957	CP003603	2789538 bp	59	-	+
S. aureus	SA40	CP003604	2728308 bp	59	-	+
S. aureus	SA268	CP006630	2833899 bp	59	-	+
S. aureus	TMUS2126	AP014652	2770164 bp	72	-	+
S. aureus	TMUS2134	AP014653	2770164 bp	72	-	+
S. aureus	CN1	CP003979	2751266 bp	72	-	+
S. aureus	11819-97	CP003194	2846546 bp	80	+	+
S. aureus	NCTC13435	LN831036	2797452 bp	80	+	+
S. aureus	JKD6159	CP002114	2811435 bp	93	+	+
S. aureus	JH9	CP000703	2906700 bp	105	-	+
S. aureus	JH1	CP000736	2906507 bp	105	-	+
S. aureus	FCFHV36	CP011147	2849811 bp	105	-	+
S. aureus	93b_S9	CP010952	2788353 bp	121	+	+
S. aureus	ED133	CP001996	2832478 bp	133	-	-
S. aureus	RF122	AJ938182	2742531 bp	151	+	+
S. aureus	SA17_S6	CP010941	2672185 bp	152	-	+
S. aureus	04-02981	CP001844	2821452 bp	225	-	+
S. aureus	10388	HE579059	2759510 bp	228	-	+
S. aureus	10497	HE579061	2759512 bp	228	-	+
S. aureus	15532	HE579063	2759883 bp	228	-	+
S. aureus	16035	HE579065	2759835 bp	228	-	+
S. aureus	16125	HE579067	2759457 bp	228	-	+
S. aureus	18341	HE579069	2759473 bp	228	-	+
S. aureus	18412	HE579071	2759263 bp	228	-	+
S. aureus	18583	HE579073	2759328 bp	228	-	+
S. aureus	JKD6008	CP002120	2924344 bp	239	+	+
S. aureus	T0131	CP002643	2913900 bp	239	+	+
S. aureus	Bmb9393	CP005288	2980548 bp	239	+	+
S. aureus	Z172	CP006838	2987966 bp	239	+	+
S. aureus	XN108	CP007447	3052055 bp	239	+	+

S. aureus	Gv69	CP009681	3046210 bp	239	+	+
S. aureus	TW20	FN433596	3043210 bp	239	+	+
S. aureus	ATCC 25923	CP009361	2778854 bp	243	+	+
S. aureus	Col	CP000046	2809422 bp	250	+	+
S. aureus	NRS 100	CP007539	2823087 bp	250	+	+
S. aureus	Newman	AP009351	2878897 bp	254	+	+
S. aureus	PS187	ARPA00000000	2781079 bp	395	-	-
S. aureus	SO385	AM990992	2872582 bp	398	-	+
S. aureus	71193	CP003045	2715000 bp	398	-	+
S. aureus	08BA02176	CP003808	2782313 bp	398	-	+
S. aureus	LGA251	FR821779	2750834 bp	425	-	+
S. aureus	M0831	KB821688	2957781 bp	609	+	+
S. aureus	DAR4145	CP010526	2860508 bp	772	-	+
S. aureus	144_S7	CP010943	2730860 bp	772	-	+
S. aureus	79_S10	CP010944	2726524 bp	772	-	+
S. aureus	SASCBU26	CDLR00000000	2862578 bp	2371	-	+
S. aureus	TCH60	CP002110	2802675 bp	*	+	+
S. aureus	M1216	KB822075	2793375 bp	**	-	+
S. argenteus	MSHR1132	FR821777	2762785 bp	1850	+	+
S. schweitzeri	FSA084	CCEL00000000	2748405 bp	2022	+	+

References

1. Thomas, C. M. & Nielsen, K. M. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature reviews.Microbiology* **3,** 711–721, doi: 10.1038/nrmicro1234 (2005).

2. Popa, O. & Dagan, T. Trends and barriers to lateral gene transfer in prokaryotes. *Current opinion in microbiology* **14,** 615–623, doi: 10.1016/j.mib.2011.07.027 (2011).

3. Barrangou, R. & Marraffini, L. A. CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Molecular cell* **54,** 234–244, doi: 10.1016/j.molcel.2014.03.011 (2014).

4. Corvaglia, A. R. *et al.* A type III-like restriction endonuclease functions as a major barrier to horizontal gene transfer in clinical Staphylococcus aureus strains. *Proceedings of the National Academy of Sciences of the United States of America* **107,** 11954–11958, doi: 10.1073/pnas.1000489107 (2010).

5. Waldron, D. E. & Lindsay, J. A. Sau1: a novel lineage-specific type I restriction-modification system that blocks horizontal gene transfer into Staphylococcus aureus and between S. aureus isolates of different lineages. *Journal of bacteriology* **188,** 5578–5585, doi: 10.1128/JB.00418-06 (2006).

6. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nature reviews. Microbiology* **8,** 317–327, doi: 10.1038/nrmicro2315 (2010).

7. Lindsay, J. A. Genomic variation and evolution of Staphylococcus aureus. *International journal of medical microbiology : IJMM* **300,** 98–103, doi: 10.1016/j.ijmm.2009.08.013 (2010).

8. Brussow, H., Canchaya, C. & Hardt, W. D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiology and molecular biology reviews: MMBR* **68,** 560–602, table of contents, doi: 10.1128/MMBR.68.3.560-602.2004 (2004).

9. Moon, B. Y. *et al.* Phage-mediated horizontal transfer of a Staphylococcus aureus virulence-associated genomic island. *Scientific reports* **5,** 9784, doi: 10.1038/srep09784 (2015).

10. Malachowa, N. & DeLeo, F. R. Mobile genetic elements of Staphylococcus aureus. *Cellular and molecular life sciences: CMLS* **67,** 3057–3071, doi: 10.1007/s00018-010-0389-4 (2010).

11. Winstel, V. *et al.* Wall teichoic acid structure governs horizontal gene transfer between major bacterial pathogens. *Nat Commun***4,** 2345, doi: 10.1038/ncomms3345 (2013).

12. Winstel, V., Sanchez-Carballo, P., Holst, O., Xia, G. & Peschel, A. Biosynthesis of the unique wall teichoic acid of Staphylococcus aureus lineage ST395. *mBio* **5,** e00869–00814, doi: 10.1128/mBio.00869-14 (2014).

13. Brown, S., Santa Maria, J. P., Jr. & Walker, S. Wall teichoic acids of gram-positive bacteria. *Annual review of microbiology* **67,** 313–336, doi: 10.1146/annurev-micro-092412-155620 (2013).

14. Weidenmaier, C. & Peschel, A. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nature reviews. Microbiology* **6,** 276–287, doi: 10.1038/nrmicro1861 (2008).

15. Xia, G. *et al.* Glycosylation of wall teichoic acid in Staphylococcus aureus by TarM. *The Journal of biological chemistry* **285,** 13405–13415, doi: 10.1074/jbc.M109.096172 (2010).

16. Brown, S. *et al.* Methicillin resistance in Staphylococcus aureus requires glycosylated wall teichoic acids. *Proceedings of the National Academy of Sciences of the United States of America* **109,** 18909–18914, doi: 10.1073/pnas.1209126109 (2012).

17. Xia, G. *et al.* Wall teichoic Acid-dependent adsorption of staphylococcal siphovirus and myovirus. *Journal of bacteriology* **193,** 4006–4009, doi: 10.1128/JB.01412-10 (2011).

18. Yang, S. *et al.* Analysis of the features of 45 identified CRISPR loci in 32 Staphylococcus aureus. *Biochemical and biophysical research communications* **464,** 894–900, doi: 10.1016/j.bbrc.2015.07.062 (2015).

19. Pantucek, R. *et al.* The polyvalent staphylococcal phage phi 812: its host-range mutants and related phages. *Virology* **246,** 241–252, doi: 10.1006/viro.1998.9203 (1998).

20. Vybiral, D. *et al.* Complete nucleotide sequence and molecular characterization of two lytic Staphylococcus aureus phages: 44AHJD and P68. *FEMS microbiology letters* **219,** 275–283 (2003).

21. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of molecular biology* **215,** 403–410, doi: 10.1016/S0022-2836(05)80360-2 (1990).

22. Winstel, V., Xia, G. & Peschel, A. Pathways and roles of wall teichoic acid glycosylation in Staphylococcus aureus. *International journal of medical microbiology: IJMM* **304,** 215–221, doi: 10.1016/j.ijmm.2013.10.009 (2014).

23. Tong, S. Y. *et al.* Novel staphylococcal species that form part of a Staphylococcus aureus-related complex: the non-pigmented Staphylococcus argenteus sp. nov. and the non-human primate-associated Staphylococcus schweitzeri sp. nov. *Int J Syst Evol Microbiol* **65,** 15–22, doi: 10.1099/ijs.0.062752-0 (2015).

24. Winstel, V. *et al.* Wall Teichoic Acid Glycosylation Governs Staphylococcus aureus Nasal Colonization. *mBio* **6,** doi: 10.1128/ mBio.00632-15 (2015).

25. Kurokawa, K. *et al.* Glycoepitopes of staphylococcal wall teichoic acid govern complement-mediated opsonophagocytosis via human serum antibody and mannose-binding lectin. *The Journal of biological chemistry* **288,** 30956–30968, doi: 10.1074/jbc. M113.509893 (2013).

26. Avrani, S., Wurtzel, O., Sharon, I., Sorek, R. & Lindell, D. Genomic island variability facilitates Prochlorococcus-virus coexistence. *Nature* **474,** 604–608, doi: 10.1038/nature10172 (2011).

27. Qimron, U., Marintcheva, B., Tabor, S. & Richardson, C. C. Genomewide screens for Escherichia coli genes affecting growth of T7 bacteriophage. *Proceedings of the National Academy of Sciences of the United States of America* **103,** 19039–19044, doi: 10.1073/ pnas.0609428103 (2006).

28. Wilkinson, B. J. & Holmes, K. M. Staphylococcus aureus cell surface: capsule as a barrier to bacteriophage adsorption. *Infection and immunity* **23,** 549–552 (1979).

29. Bondy-Denomy, J., Pawluk, A., Maxwell, K. L. & Davidson, A. R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493,** 429–432, doi: 10.1038/nature11723 (2013).

30. Bae, T. & Schneewind, O. Allelic replacement in Staphylococcus aureus with inducible counter-selection. *Plasmid* **55,** 58–63, doi: 10.1016/j.plasmid.2005.05.005 (2006).

31. Monk, I. R., Shah, I. M., Xu, M., Tan, M. W. & Foster, T. J. Transforming the untransformable: application of direct transformation to manipulate genetically Staphylococcus aureus and Staphylococcus epidermidis. *mBio* **3,** e00277–00211 doi: 10.1128/ mBio.00277-11 (2012).

32. Bruckner, R. A series of shuttle vectors for Bacillus subtilis and Escherichia coli. *Gene* **122,** 187–192 (1992).

33. Enright, M. C., Day, N. P., Davies, C. E., Peacock, S. J. & Spratt, B. G. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of Staphylococcus aureus. *Journal of clinical microbiology* **38,** 1008–1015 (2000).

34. Winstel, V., Kuhner, P., Krismer, B., Peschel, A. & Rohde, H. Transfer of plasmid DNA to clinical coagulase-negative staphylococcal pathogens by using a unique bacteriophage. *Applied and environmental microbiology* **81,** 2481–2488, doi: 10.1128/AEM.04190-14 (2015).

35. Kurtz, S. *et al.* Versatile and open software for comparing large genomes. *Genome biology* **5,** R12, doi: 10.1186/gb-2004-5-2-r12 (2004).

36. Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Molecular biology and evolution* **23,** 254–267, doi: 10.1093/molbev/msj030 (2006).

37. Kreiswirth, B. N. *et al.* The toxic shock syndrome exotoxin structural gene is not detectably 134 transmitted by a prophage. *Nature* **305**, 709-712 (1983).

38. Fitzgerald, J. R., Hartigan, P. J., Meaney, W. J. & Smyth, C. J. Molecular population and virulence factor analysis of Staphylococcus aureus from bovine intramammary infection. *Journal of applied microbiology* **88**, 1028-1037 (2000).

39. Ben Zakour, N. L. *et al.* Genome-wide analysis of ruminant Staphylococcus aureus reveals diversification of the core genome. *Journal of bacteriology* **190**, 6302-6317, doi:10.1128/JB.01984-07 (2008).

40. Dyke, K. G., Jevons, M. P. & Parker, M. T. Penicillinase production and intrinsic resistance to penicillins in Staphylococcus aures. *Lancet* **1**, 835-838 (1966).

41. Asheshov, E. A. & Jevons, M. P. The effect of heat on the ability of a host strain to support 144 the growth of a Staphylococcus phage. *Journal of general microbiology* **31**, 97-107 (1963).

42. Centers for Disease, C. & Prevention. Four pediatric deaths from community-acquired methicillin-resistant Staphylococcus aureus - Minnesota and North Dakota, 1997-1999. *MMWR. Morbidity and mortality weekly report* **48**, 707-710 (1999).

43. Kuroda, M. *et al.* Whole genome sequencing of meticillin-resistant Staphylococcus aureus. *Lancet* **357**, 1225-1240 (2001).

44. Mwangi, M. M. *et al.* Tracking the in vivo evolution of multidrug resistance in Staphylococcus aureus by whole-genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 9451-9456, doi:10.1073/pnas.0609839104 (2007).

45. Li, M. *et al.* MRSA epidemic linked to a quickly spreading colonization and virulence determinant. *Nature medicine* **18**, 816-819, doi:10.1038/nm.2692 (2012).

46. Schleifer, K. H. & Kloos, W. E. Isolation and Characterization of Staphylococci from Human Skin .1. Amended Descriptions of Staphylococcus-Epidermidis and Staphylococcus- Saprophyticus and Descriptions of 3 New Species - Staphylococcus-Cohnii, Staphylococcus- Haemolyticus, and Staphylococcus-Xylosus. *International Journal of Systematic Bacteriology* **25**, 50-61 (1975).

47. Mack, D., Siemssen, N. & Laufs, R. Parallel induction by glucose of adherence and a polysaccharide antigen specific for plastic-adherent Staphylococcus epidermidis: evidence for functional relation to intercellular adhesion. *Infection and immunity* **60**, 2048-2057 (1992).

48. O'Flaherty, S. *et al.* Potential of the polyvalent anti-Staphylococcus bacteriophage K for control of antibiotic-resistant staphylococci from hospitals. *Applied and environmental microbiology* **71**, 1836-1842, doi:10.1128/AEM.71.4.1836-1842.2005 (2005).

49. Goerke, C. *et al.* Role of Staphylococcus aureus global regulators sae and sigmaB in virulence gene expression during device-related infection. *Infection and immunity* **73**, 3415-3421, doi:10.1128/IAI.73.6.3415-3421.2005 (2005).

Chapter 5

-

General discussion

Strategies for affecting horizontal gene transfer (HGT) in different staphylococcal hosts

HGT is frequently occurring in Staphylococcal species

Staphylococcal species are important bacteria both as commensals and pathogens closely related to humans and animals. The genus Staphylococcus currently contains more than 47 species and 23 subspecies, which can be grouped into 15 clusters based on a phylogenetic analysis performed using DNA sequence data from multiple loci, such as the 16S rRNA gene and *dnaJ*, *rpoB* and *tuf* gene fragments[1,2]. Among all the Staphylococcus species, coagulase-negative *staphylococci* (CoNS), not only *S. epidermidis*, are gaining increasing interest nowadays for their increasing roles in nosocomial infections[3,4,5]. However, it remains largely unknown how *S. epidermidis* and other Staphylococcal species switch from commensal to pathogenic and how the Staphylococcal family becomes resistant to more and more antibiotics.

A study of bacterial evolution estimated that approximately 20% of the gene content of bacteria is considered to be obtained from other bacterial strains within the same or different species[6]. In the evolution of the bacterial population, horizontal gene transfer plays an important role in bacterial pathogenesis by helping bacteria acquire antibiotic resistance genes, adapt to the environment and proliferate in host tissues. The elements involved in horizontal gene transfer include temperate phages, plasmids, transposons and other horizontally acquired units. Horizontal transfer of phages is an efficient way to rapidly disseminate virulence determinants among pathogens. With the increasing number of phage-host interaction studies and the development of genome sequencing data, the recent knowledge of *S. epidermidis* and their phages has contributed to our understanding of evolution, the emergence of virulence strains, and the co-survival of Staphylococcal species.

HGT could promote colonization, antibiotic resistance, and virulence in Staphylococci

The arginine catabolic mobile element (ACME) is a nice example of a factor obtained by the most important Staphylococcal pathogen-*S. aureus*. ACME is a mobile genetic element and approximately 31 kb. It was first reported in the community-associated methicillin-resistant *Staphylococcus aureus* (CA-MRSA) strain USA300 located just beside the type IV SCC*mec* element[7]. Until now, three types of ACMEs have been found in *S. aureus*. ACME type I was the first to be reported and contains the gene *opp3* (an oligopeptide permease) and gene *arc* (an arginine deiminase); type II contains only *arc*; and type III contains only *opp3*. Notably, ACME, especially USA300-ACME type I, is very frequently found in the genome of *S. epidermidis*[8]. This finding indicates that the ACME elements found in *S. aureus* originated from *S. epidermidis*. In a study by Joshi *et al.*, the ACME element contains the gene *speG*, which helps USA300 strains colonize human skin by withstanding levels of polyamines such as spermidine that are produced in skin and toxic to other *S. aureus*[9].

Methicillin resistance is encoded by the *mecA* gene, which encodes a penicillin resistance gene with decreased binding affinity to methicillin[10]. The *mecA* gene is located on a mobile genetic element named staphylococcal cassette chromosome *mec* (SCC*mec*) in s*taphylococci*[11]. There is strong evidence that the SCC*mec* element in *S. aureus* originated from *S. epidermidis*. Type IV SCC*mec* of *S. aureus* and *S. epidermidis* have high homology of more than 98%[12]. Methicillin resistance is also much higher in *S. epidermidis* than in *S. aureus*[13]. The mechanism of SCC*mec* transfer is one of the most important questions that has remained a mystery until now. All transformation methods, including conjugation, were tested in laboratory experiments but failed[14]. Phage-mediated transduction of SCC*mec* elements was successful between *S. aureus* strains but not interspecies between other Staphylococcal species[15]. The recently discovered *S. epidermidis* with two types of WTA might be helpful in elucidating this mechanism because this strain shares similar WTA structures with *S. aureus* that may be recognized by the same phages for possible transduction (see Chapter 2).

This newly found *S. epidermidis* with two types of WTA could also be an example for bacteria enhancing their invasiveness by the HGT process (see Chapter 2). Compared to the common *S. epidermidis* strains, the *S. epidermidis* with two types of WTA was only prevalent in infection strains from the hospital but not in any of the strains isolated from healthy people. The additional novel *S. aureus*-like WTA structure, encoded by *the tarIJLM* gene cluster, was suspected to be an important factor in bacterial virulence. The wild-type strain is more virulent in animal models compared to its *tarIJLM2* knockout strain, surviving better and causing more death in the bacteremia model. The evolutionary tree showed that the *tarIJL2* gene in this novel *tarIJLM2* gene cluster was more similar to *S. aureus tarIJL* than the conserved *tarIJL1* in *S. epidermidis*. The novel *tarM2* gene was also found to be similar to the

sugar transferase gene *tarM* in *S. aureus*. These results indicated that the virulence-related *tarIJLM2* gene might have been shared between *S. epidermidis* and *S. aureus* at one time. The existence of *tarIJLM2* in different lineages in the evolutionary tree also indicated that multiple HGT events occurred in this element.

Strategies for affecting HGT from the side of the phage

I. Structure variation in the phage tail

Interspecies gene exchanges might be of great importance for the pathogenesis of Staphylococcal species. The host recognition specificity of unrelated phages can be altered via horizontal gene transfer of the tail fibre genes, which are likely involved in host-phage interactions, among unrelated phages. An example of this is the putative tail-associated Zn carboxypeptidase in the *S. hominis* phages StB12 and StB27[16]. This tends to occur under environmental selective pressure[17]. Phages could use this strategy to alter their specificity towards the bacterial host. This might occur independently without changing other structural components of the phage. Several studies have shown that host range shifts in phages are efficiently improved by phage-bacterium coevolution[18,19].

For *S. aureus* phage Φ11, the receptor-binding protein was reported in 2016[20]. Gp45 is a tail protein localized in the baseplate of Φ11. The siphovirus Φ11 can bind to both α- and β-GlcNAc residues on WTA. The knockout of both *tarM* and *tarS* genes can lead to a significant reduction in the adsorption of Φ11 to the *S. aureus* mutant strain. The elucidation of Gp45-involved molecular interactions helped to better understand siphovirus-mediated HGT. The protein structure of Gp45 has also been reported. Electron microscopy shows that six receptor-binding protein (RBP) trimers are assembled around the baseplate core. Each monomer contains a five-bladed propeller domain with a cavity that could accommodate a GlcNAc moiety. This structure is conserved among most glycan-recognizing *Siphoviridae*[21].

II. The variation of integrase

Compared to *S. aureus* prophages, the integrate sites were quite diverse and most of them were reported to insert into the β-haemolysin (hlb) or lipase genes (geh)[22] (Table 1). This characteristic of integrating site diversity might correspond to the diversity of *S. epidermidis* species. Conversely, this could also contribute to promoting the exchange of genes between different strains by increasing the infection frequencies by *S. epidermidis* phages. Analysis of several *S. epidermidis* prophages with some of the integrase-type defined prophages from *S. aureus* revealed that the integrases of *S. epidermidis* prophages and other CoNS prophages were closely related to those from *S. aureus* in the integrase-based evolutionary tree (Figure 1. Du *et al*. unpublished data). Similar to what has been observed in the genome of *S. aureus*, most strains of *S. epidermidis* have more than one prophage in their genomes. In addition, some prophages in the genome of *S. epidermidis* could have more than one integrase. However, the reason is unknown.

Table 1. Integrating sites of prophages in chromosome of *S. epidermidis* strains.

Host strain	Location	Size (kb)	Integrating site
1457	1794046-1834721	40.6	tRNA-Ser
RP62A	1563904-1697887	133.9	hypothetical protein
DAR1907	612279-655144	42.8	tRNA-Ser
DAR1907	1680737-1823843	143.1	LacI family transcriptional regulator
HD33	1711803-1758453	46.6	tRNA-Ser
NCTC13924	1007166-1042551	35.3	glnA type I glutamate-ammonia ligase
NCTC13924	1777588-1925438	147.8	YeeE/YedE family protein
NCTC13924	2385015-2427981	42.9	resolvase, N terminal domain protein
14.1.R1	1190601-1249965	59.3	SAM-dependent methyltransferase
14.1.R1	1901793-1956546	44.7	pyridine nucleotide-disulfide oxidoreductase family protein

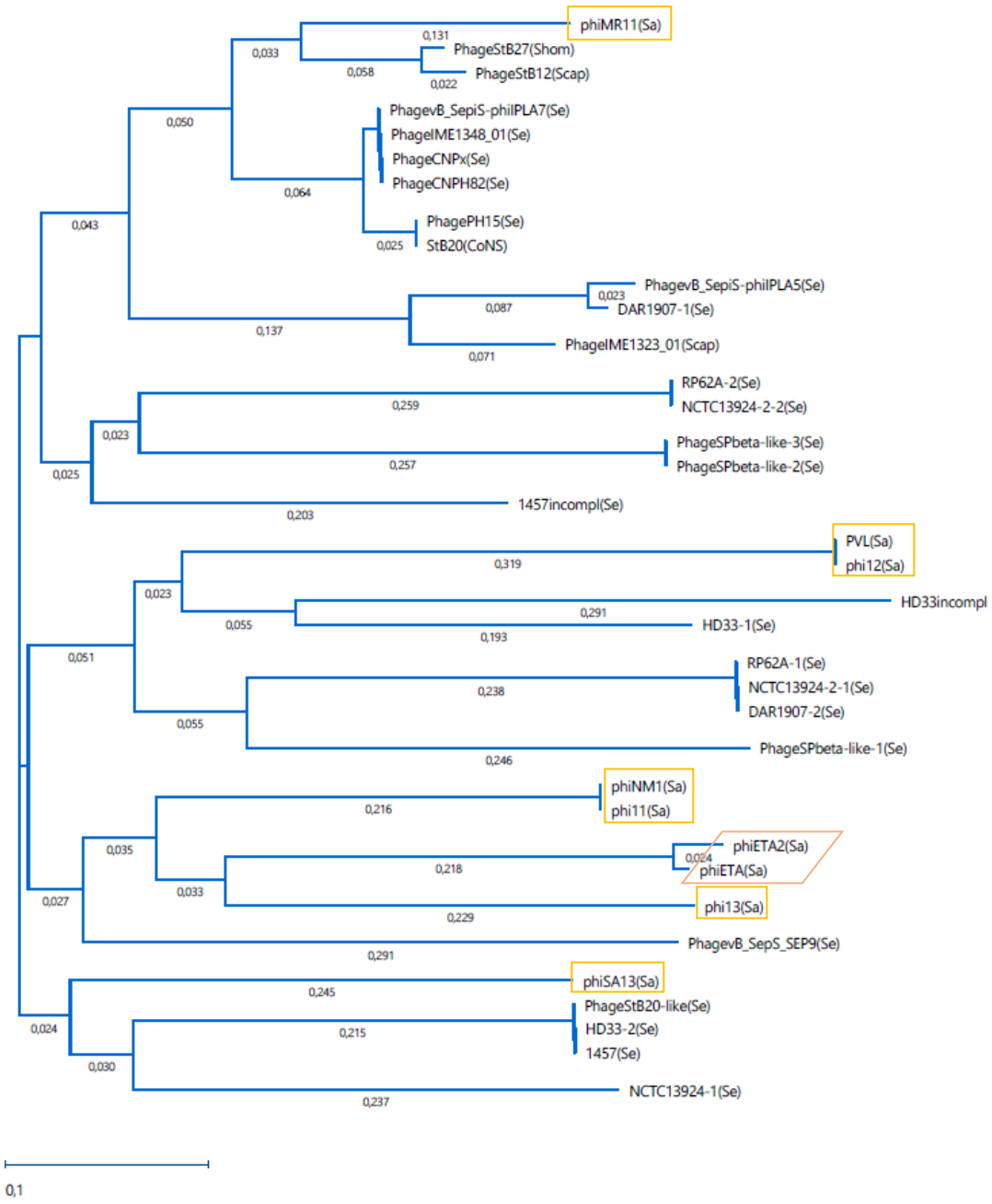


Figure 1. Phylogenetic tree of staphylococcal bacteriophages in the *Siphoviridae* family. (Sa), Prophage from *S. aureus*; (Se), prophage from *S. epidermidis*; (CoNS), prophage from CoNS; (Shom), prophage from *S. hominis*; (Scap), prophage from *S. capitis*. The yellow boxes mark the integrase group of *S. aureus*. The integrase clusters of *S. aureus* are marked with orange squares (from up to down): cluster Sa12, cluster Sa2, cluster Sa5, cluster Sa1, cluster Sa3, and phiSA13 are of unknown clusters.

Strategies for affecting HGT from the side of the bacteria

I. Structure variation in wall teichoic acid

Wall teichoic acid (WTA) is the most abundant molecule on the cell wall of s*taphylococci*. Phages can adsorb to the wall teichoic acid of bacteria as a specific receptor for further infection and replication. Different Staphylococcal species have different WTA structures that can be recognized by different phages. *S. epidermidis* and other CoNS have similar WTA structures composed of polymers of 1,3-glycerol-phosphate (GroP) with α/β-glucose or α-N-acetylglucosamine (GlcNAc)[23] [24](see Chapter 2). *S. capitis* and *S. hominis* have GroP polymers decorated with either α- or β-GlcNAc or both[60]. The coagulase-positive Staphylococcus *S. aureus* usually has WTA composed of 1,5-ribitol-phosphate (RboP) polymers decorated with either α- or β-GlcNAc or both[25,26]. Despite the different sugar decorations observed in different Staphylococcal species and strains, all of the WTA polymers reported from the Staphylococcal family are modified with alanyl groups.

The distinct structures of WTAs are the main barrier for HGT via phage transduction for different Staphylococcal species. Recently, strains of certain sequence types from *S. epidermidis* and *S. aureus* were reported to present special WTA structures. *S. aureus* ST395 presents a CoNS-similar WTA structure consisting of GroP-polymer and α-GalNAc[27]. Recently, we found some special sequence types (ST) (ST10, ST23, and ST87) of *S. epidermidis* presenting an additional *S. aureus* type WTA with RboP-polymers decorated with α-glucose on their surfaces together with the *S. epidermidis*-common GroP-polymer WTA structure with no sugar decoration (see Chapter 2). This *S. epidermidis* with two types of WTA could be infected by both *S. epidermidis* phages and *S. aureus* phages and thus could act as a 'bridge' for phage transduction in interspecies HGT (Figure 2) (see Chapter 2). This newly found evidence of similar WTA shared between different Staphylococcus species supports the hypothesis that s*taphylococci* can exchange genes between them.

Li *et al.* at Tübingen University reported that while TarS-mediated WTA β-O-GlcNAcylation is required for the susceptibility of *S. aureus* to podoviruses, TarM-mediated WTA α-O-GlcNAcylation protects *S. aureus* from infection by short-tailed, lytic *Podoviridae*[28]. The evolutionary study showed that *tarM* is not contained in every *S. aureus* strain. Some phylogenetic branches lost the *tarM* gene during their evolution, leading to the susceptibility of *S. aureus* to podovirus infection[28]. The reason why these *S. aureus* strains lost the *tarM* gene is still waiting to be elucidated. However, the *S. aureus* strains containing *tarM* can survive in the environment better than the strains lacking *tarM*. The *tarS* gene, which has been detected in all *S. aureus* strains, can ensure the evolution of *S. aureus* strains containing *tarM*, allowing these strains to undergo the process of evolution with other s*taphylococci*, such as these *S. epidermidis* strains with two types of WTA.

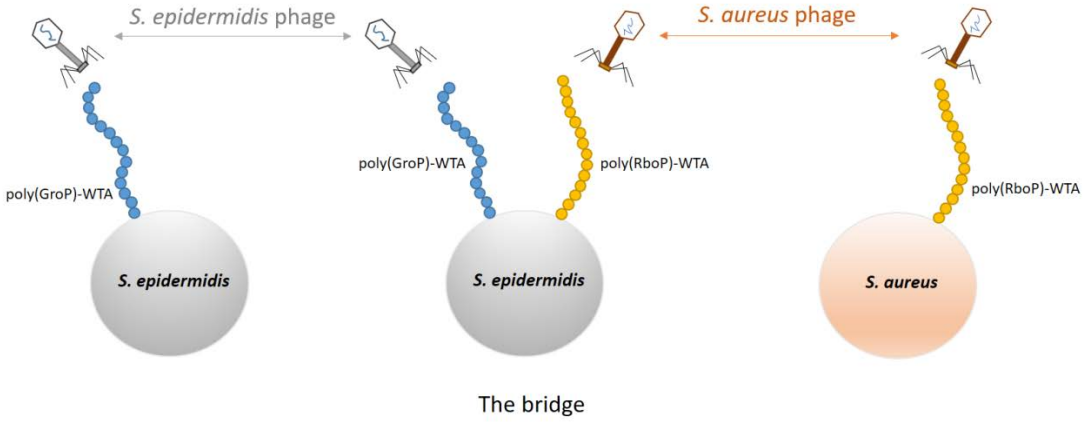


Figure 2. Schematic of the novel *S. epidermidis* with two types of WTA acting as a HGT bridge between *S. aureus* and *S. epidermidis* (see Chapter 2).

R-M systems and CRISPR direct and balance HGT in *staphylococci*

Bacteria are always challenged by phage attacks that threaten their successful survival in the environment. To prevent phage infection, bacteria have built up restriction-modification (R-M) systems that degrade invading DNA[29,30]. R-M systems function in two steps. First, methyltransferases methylate adenine or cytosine bases at certain recognition sites. Then, restriction endonucleases recognize and catalyse double strand cleavage of the same sequence that has not been modified[30.31]. R-M systems do not destroy the host DNA since the host DNA is protected by methylation. By recognition sequences, cut sites, cleavage specifications and structures, R-M systems are classified into four types: I, II, III, and IV. A type II restriction endonuclease, SepII, from *S. epidermidis* was reported to have the ability to cleave DNA effectively, even maintaining full activity in 100 mM NaCl, which is considered a strong barrier for HGT events[32].

Clustered regularly interspaced short palindromic repeats (CRISPR) is another important system in bacteria that fights against non-self DNA from the environment. The CRISPR system contains a CRISPR-associated (cas) protein, spacers and repeats. The spacers are transcribed and processed into small CRISPR RNAs that guide the cas protein to foreign DNA, which is recognized and destroyed. More than half of the *S. epidermidis* strains contain CRISPR sequences, while much fewer *S. aureus* strain contain one. CRISPR might play a role in preventing the uptake of foreign DNA in *S. epidermidis*. This is also strong evidence that CRISPR might direct the gene flow from *S. epidermidis* to *S. aureus*[33]. In our study (see Chapter 2) of the *S. epidermidis* with two types of WTA, CRISPR-cas 6 was found frequently in ST5, ST10, and ST23

S. epidermidis, which were the strains with a high rate of having two types of WTA (Table 2). The *S. epidermidis* with two types of WTA tended to obtain foreign DNA much easier with more phage receptors on the surface than other *S. epidermidis* strains. These *S. epidermidis* strains might use the CRISPR system to balance the prevention and acquisition of harmful genes from other bacteria to maintain survival.

Table 2. CRISPR (cas6) is found frequently only in ST5, ST10, and ST23 *S. epidermidis*, which have a high rate of two types of WTA.

S. epidermidis ST types	CRISPR(cas6) positive (positive/total number, percentage)
2	0
5	8/15, 53.3%
7	0
8	0
10	1/4, 25.0%
16	0
20	0
21	0
22	0
23	30/60, 50.0%
SLV of ST23	3/3, 100%
46	0
59	0
86	0
87	0
89	0
100	0
143	0
184	0
188	0
210	0
218	0
225	0
368	0
487	0

Conclusion and perspectives

More and more studies suggest that different Staphylococcus species exchange genes via HGT to promote successful survival in various environmental conditions. With frequent acquisition and gene loss, *S. epidermidis* seems to play an important role as a ‘transportation station‘, transporting genes from *S. aureus* to CoNS and from CoNS to *S. aureus*. This also leads to *S. epidermidis* becoming a more diverse bacterial species. The discoveries of ST395 *S. aureus* expressing CoNS-like WTA and *S. epidermidis* with two types of WTA with the *S. aureus*-like WTA structure provide researchers with more significant evidence. With the emergence of the *S. epidermidis* with two types of WTA, the virulence genes might spread among more strains and provide a possibility for new virulent strains to infect both humans and animals. CRISPR together with the R-M system protected these easier-to-transduce *S. epidermidis* from unnecessary and harmful foreign genes. However, there are remaining questions. Hopefully, with the increasing genomic data from both bacteria and phages and the illustration of wall teichoic acid or other surface structures of s*taphylococci*, together with the investigation of the biosynthesis pathway, researchers will be able to unveil the mysteries of gene evolution in *S. epidermidis* and other s*taphylococci* and their interactions.

References

1 Becker, K., Heilmann, C. & Peters, G. Coagulase-Negative Staphylococci. *Clin Microbiol Rev* **27**, 870-926, doi:10.1128/Cmr.00109-13 (2014).

2 Lamers, R. P. *et al.* Phylogenetic relationships among Staphylococcus species and refinement of cluster groups based on multilocus data. *Bmc Evol Biol* **12**, doi:Artn 17110.1186/1471-2148-12-171 (2012).

3 Butin, M., Martins-Simoes, P., Rasigade, J. P., Picaud, J. C. & Laurent, F. Worldwide Endemicity of a Multidrug-Resistant Staphylococcus capitis Clone Involved in Neonatal Sepsis. *Emerg Infect Dis* **23**, 538-539, doi:10.3201/eid2303.160833 (2017).

4 Simoes, P. M. *et al.* Single-Molecule Sequencing (PacBio) of the Staphylococcus capitis NRCS-A Clone Reveals the Basis of Multidrug Resistance and Adaptation to the Neonatal Intensive Care Unit Environment. *Front Microbiol* **7**, 1991, doi:10.3389/fmicb.2016.01991 (2016).

5 Al Hennawi, H. E. T., Mahdi, E. M. & Memish, Z. A. Native valve Staphylococcus capitis infective endocarditis: a mini review. *Infection*, doi:10.1007/s15010-019-01311-8 (2019).

6 Gogarten, J. P., Doolittle, W. F. & Lawrence, J. G. Prokaryotic evolution in light of gene transfer. *Mol Biol Evol* **19**, 2226-2238, doi:DOI 10.1093/oxfordjournals.molbev.a004046 (2002).

7 Diep, B. A. *et al.* Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant Staphylococcus aureus. *Lancet* **367**, 731-739, doi:10.1016/S0140-6736(06)68231-7 (2006).

8 Barbier, F. *et al.* High prevalence of the arginine catabolic mobile element in carriage isolates of methicillin-resistant Staphylococcus epidermidis. *J Antimicrob Chemother* **66**, 29-36, doi:10.1093/jac/dkq410 (2011).

9 Joshi, G. S., Spontak, J. S., Klapper, D. G. & Richardson, A. R. Arginine catabolic mobile element encoded speG abrogates the unique hypersensitivity of Staphylococcus aureus to exogenous polyamines. *Mol Microbiol* **82**, 9-20, doi:10.1111/j.1365-2958.2011.07809.x (2011).

10 Chambers, H. F., Hartman, B. J. & Tomasz, A. Increased amounts of a novel penicillin-binding protein in a strain of methicillin-resistant Staphylococcus aureus exposed to nafcillin. J Clin Invest 76, 325-331, doi:10.1172/JCI111965 (1985).

11 Katayama, Y., Ito, T. & Hiramatsu, K. A new class of genetic element, staphylococcus cassette chromosome mec, encodes methicillin resistance in Staphylococcus aureus. *Antimicrob Agents Chemother* **44**, 1549-1555, doi:10.1128/aac.44.6.1549-1555.2000 (2000).

12 Jones, R. N., Barry, A. L., Gardiner, R. V. & Packer, R. R. The prevalence of staphylococcal resistance to penicillinase-resistant penicillins. A retrospective and prospective national surveillance trial of isolates from 40 medical centers. *Diagn Microbiol Infect Dis* **12**, 385-394, doi:10.1016/0732-8893(89)90108-9 (1989).

13 Schmitz, F. J., Verhoef, J. & Fluit, A. C. Prevalence of resistance to MLS antibiotics in 20 European university hospitals participating in the European SENTRY surveillance programme. Sentry Participants Group. *J Antimicrob Chemother* **43**, 783-792, doi:10.1093/jac/43.6.783 (1999).

14 Lacey, R. W. Transfer of chromosomal genes between staphylococci in mixed cultures. *J Gen Microbiol* **71**, 399-401, doi:10.1099/00221287-71-2-399 (1972).

15 Cohen, S. & Sweeney, H. M. Effect of the prophage and penicillinase plasmid of the recipient strain upon the transduction and the stability of methicillin resistance in Staphylococcus aureus. *J Bacteriol* **116**, 803-811 (1973).

16 Deghorain, M. *et al.* Characterization of Novel Phages Isolated in Coagulase-Negative Staphylococci Reveals Evolutionary Relationships with Staphylococcus aureus Phages. *Journal of Bacteriology* **194**, 5829-5839, doi:10.1128/Jb.01085-12 (2012).

17 Haggardljungquist, E., Halling, C. & Calendar, R. DNA-Sequences of the Tail Fiber Genes of Bacteriophage P2 - Evidence for Horizontal Transfer of Tail Fiber Genes among Unrelated Bacteriophages. *Journal of Bacteriology* **174**, 1462-1477, doi:DOI 10.1128/jb.174.5.1462-1477.1992 (1992).

18 Hall, A. R., Scanlan, P. D. & Buckling, A. Bacteria-Phage Coevolution and the Emergence of Generalist Pathogens. *Am Nat* **177**, 44-53, doi:10.1086/657441 (2011).

19 Meyer, J. R. *et al*. Repeatability and Contingency in the Evolution of a Key Innovation in Phage Lambda. Science 335, 428-432, doi:10.1126/science.1214449 (2012).

20 Li, X. *et al*. An essential role for the baseplate protein Gp45 in phage adsorption to Staphylococcus aureus. Sci Rep 6, 26455, doi:10.1038/srep26455 (2016).

21 Koc, C. *et al*. Structure of the host-recognition device of Staphylococcus aureus phage varphi11. Sci Rep 6, 27581, doi:10.1038/srep27581 (2016).

22 Goerke, C. *et al*. Diversity of prophages in dominant Staphylococcus aureus clonal lineages. J Bacteriol 191, 3462-3468, doi:10.1128/JB.01804-08 (2009).

23 Sadovskaya, I., Vinogradov, E., Li, J. J. & Jabbouri, S. Structural elucidation of the extracellular and cell-wall teichoic acids of Staphylococcus epidermidis

RP62A, a reference biofilm-positive strain. *Carbohyd Res* **339**, 1467-1473, doi:10.1016/j.carres.2004.03.017 (2004).

24 Endl, J., Seidl, P. H., Fiedler, F. & Schleifer, K. H. Determination of Cell-Wall Teichoic-Acid Structure of Staphylococci by Rapid Chemical and Serological Screening Methods. *Arch Microbiol* **137**, 272-280, doi:Doi 10.1007/Bf00414557 (1984).

25 Weidenmaier, C. & Peschel, A. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nature Reviews Microbiology* **6**, 276-287, doi:10.1038/nrmicro1861 (2008).

26 Xia, G. *et al.* Glycosylation of wall teichoic acid in Staphylococcus aureus by TarM. *J Biol Chem* **285**, 13405-13415, doi:10.1074/jbc.M109.096172 (2010).

27 Winstel, V. *et al.* Wall teichoic acid structure governs horizontal gene transfer between major bacterial pathogens. *Nat Commun* **4**, 2345, doi:10.1038/ncomms3345 (2013).

28 Li, X. *et al.* An accessory wall teichoic acid glycosyltransferase protects Staphylococcus aureus from the lytic activity of Podoviridae. *Sci Rep* **5**, 17219, doi:10.1038/srep17219 (2015).

29 Bourniquel, A. A. & Bickle, T. A. Complex restriction enzymes: NTP-driven molecular motors. *Biochimie* **84**, 1047-1059, doi:10.1016/s0300-9084(02)00020-2 (2002).

30 Pingoud, A. & Jeltsch, A. Structure and function of type II restriction endonucleases. *Nucleic Acids Res* **29**, 3705-3727, doi:10.1093/nar/29.18.3705 (2001).

31 Pingoud, A. & Jeltsch, A. Recognition and cleavage of DNA by type-II restriction endonucleases. *Eur J Biochem* **246**, 1-22, doi:10.1111/j.1432-1033.1997.t01-6-00001.x (1997).

32 Belkebir, A. & Azeddoug, H. Purification and characterization of SepII a new restriction endonuclease from Staphylococcus epidermidis. *Microbiol Res* **167**, 90-94, doi:10.1016/j.micres.2011.03.005 (2012).

33 Otto, M. Coagulase-negative staphylococci as reservoirs of genes facilitating MRSA infection: Staphylococcal commensal species such as Staphylococcus epidermidis are being recognized as important sources of genes promoting MRSA colonization and virulence. *Bioessays* **35**, 4-11, doi:10.1002/bies.201200112 (2013).

Contributions to publications or manuscripts

Chapter 2 – **Nosocomial *Staphylococcus epidermidis* remodels surface glycopolymers to shift from commensal to pathogen behavior.**

I performed all experiments but the following:
Bioinformatics (J. Larsen, M. Stegger), NMR measurements and analysis (P. Sanchez-Carballo, K.A. Duda), LC-MS measurements and analysis (A. Walter, C. Mayer), qRT-PCR (A. Both, H. Rohde), animal model (Y. Liu, J. Liu, M. Li).

Chapter 3 – **A novel *Staphylococcus epidermidis* phage ΦTÜB: the genetic characteristics and its function as a tool for high efficient plasmid transduction.**

I performed all experiments (by myself or in cooperation) but the following:
Phage genome sequencing (T. Botka, I. Maslanova), electron microscopy (P. Bardy).

Chapter 4 – **An accessory wall teichoic acid glycosyltransferase protects *Staphylococcus aureus* from the lytic activity of *Podoviridae*.**

I performed MLST analysis of PS44A, constructed the *tarM* gene complements of all the LA-MRSA strains, genomic DNA isolation, *tarS*/*M* gene screen and diluted podovirus spot assay of all the LA-MRSA strains and their *tarM* gene complement strains.

Curriculum vitae

Personal details:

Name, surname: Du, Xin

Nationality: China

Date of birth: 22.02.1987

Email: dx0057@126.com

Education:

10.2014 – ongoing

Ph.D. student at the University of Tübingen, Interfaculty Institute of Microbiology and Infection Medicine (Research Team of Prof. Dr. Andreas Peschel), Germany

09.2009 – 07.2014

Doctor of Medicine (Dr.med.) at Medical School of Fudan University, China. (Top 4 University in China)

09.2005 – 07.2009

Bachelor of Medicine (MB) at Medical School of Shanghai Jiaotong University, China. (Top 5 University in China)

Social Work:

06.05.2017-07.05.2017 attend Inno-China Entrepreneurship Competition Germany in Berlin

07.07.2016-23.07.2016 member of the Sino-German scholar team to Tibet Universities and institutes for scientific and technical support in Medical field

25.06.2016 joint in *Staphylococci* team in ‘100km-Staffellauf der Universität Tübingen‘

09.2014-pres. member of Chinese scientific forum Boyanstem in Tübingen

Scientific Publications:

1. An accessory wall teichoic acid glycosyltransferase protects Staphylococcus aureus from the lytic activity of Podoviridae.
   Li X, Gerlach D, Du X, Larsen J, Stegger M, Kühner P, Peschel A, Xia G, Winstel V.
   Scientific Reports 5 (2015): 17219.
2. Transposon Mutagenesis Identifies Novel Genes Associated with Staphylococcus aureus Persister Formation.
   Wang W, Chen J, Chen G, Du X, Cui P, Wu J, Zhao J, Wu N, Zhang W, Li M, Zhang Y.
   Frontiers in Microbiology 6 (2015): 1437.
3. Targeting surface protein SasX by active and passive vaccination to reduce Staphylococcus aureus colonization and infection.
   Liu Q*, Du X*(co-first author), Hong X, Li T, Zheng B, He L, Wang Y, Otto M, Li M.
   Infection and Immunity 83.5 (2015): 2168-2174.
4. Genetic and phenotypic characterization of Candida albicans strains isolated from infectious disease patients in Shanghai.
   Hu L, Du X, Li T, Song Y, Zai S, Hu X, Zhang X, Li M.
   Journal of Medical Microbiology. 64.1 (2015): 74-83.
5. MRSA epidemic linked to a quickly spreading colonization and virulence determinant.
   Li M*, Du X*(co-first author), Villaruz AE, Diep BA, Wang D, Song Y, Tian Y, Hu J, Yu F, Lu Y, Otto M.
   Nature Medicine. 18.5 (2012): 816.
6. Molecular Analysis of Staphylococcus epidermidis Strains Isolated from Community and Hospital Environments in China.
   Du X*, Zhu Y*(co-first author), Song Y, Li T, Luo T, Sun G, Yang C, Cao C, Lu Y, Li M. Plos One 8.5 (2013): e62742.
7. Green synthesis of silk fibroin-silver nanoparticle composites with effective antimicrobial and biofilm-disrupting properties.
   Fei X, Jia M, Du X, Yang Y, Zhang R, Shao Z, Zhao X, Chen X.
   Biomacromolecules 14.12 (2013): 4483-4488.
8. A Simple Animal Model of Staphylococcus aureus Biofilm in Sinusitis.
   Jia M, Chen Z, Du X, Guo Y, Sun T, Zhao X.
   Am J Rhinol Allergy 28.2 (2014): e115-e119.

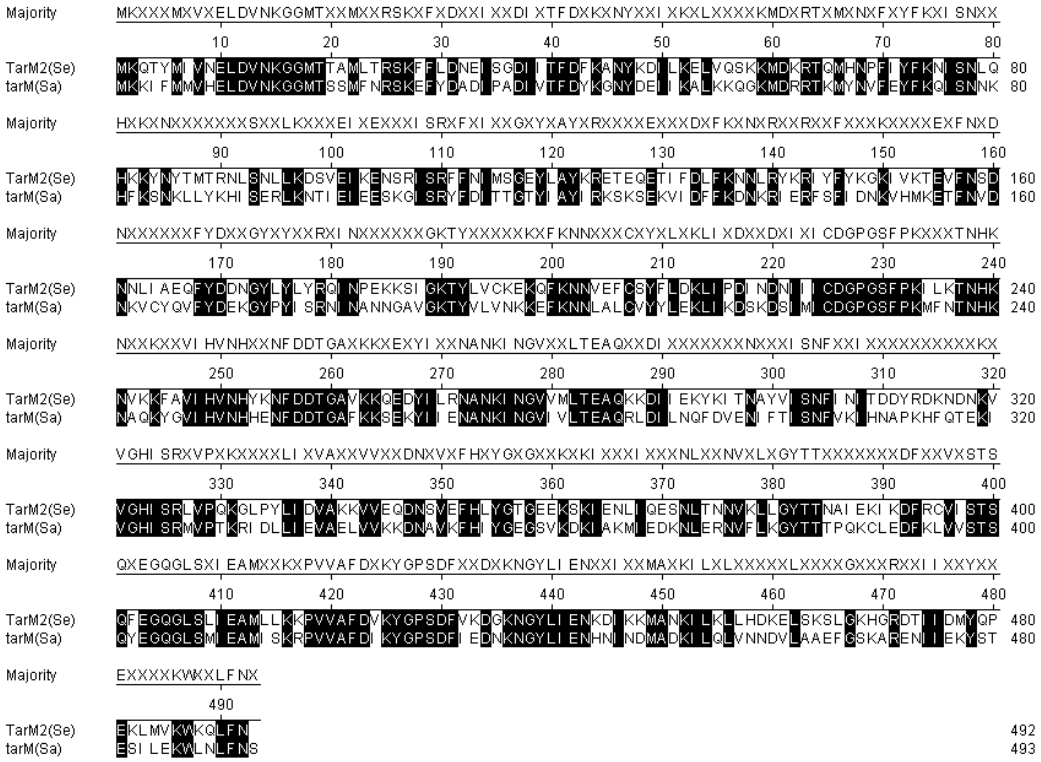
Appendix

Results and discussion for the novel *tarM2* gene

According to the result of the main Ph.D. project in this thesis (see Chapter 2), NMR indicated that the *S. epidermidis* with two types of WTA has only one kind of sugar, α-glucose, on the poly(RboP)-WTA backbone but not on the poly(GroP)-WTA backbone. The gene cluster *tarIJLM2* has been studied with clear knockout mutants and complement strains of *S. epidermidis* E73 with two types of WTA; it was determined that *tarIJLM2* is responsible for the expression of the poly(RboP)-WTA with α-glucose. A *S. epidermidis* strain 1457 with a single poly(GroP)-WTA was also transformed with a constructed plasmid, pRB474-*tarIJLM2*, for expressing the gene cluster *tarIJLM2*. This transformed strain named 1457 (pRB-*tarIJLM2*) successfully expressed additional poly(RboP)-WTA with α-glucose. From these results, it was speculated that the *tarM2* gene (previously named *tarM2*) might be the gene encoding α-glucosyltransferase.

TarM2 shares high amino acid sequence identity (51%) with the α-glycosyltransferase TarM found in *S. aureus* (Appendix Fig 1.). However, TarM2 might be the transferase for α-glucose instead of α-GlcNAc. To study this, a knockout mutant of E73 was constructed using pBASE6 and four primers (tarM_a, CGATGGTACCGCTTTATTTAAAAGAAATATATCTGATAGAAG; tarM_b, TAATATTACCTCATTATTTATTTCTTAAATGC; tarM_c, TAAATAATGAGGTAATATTAAGTATCCATTTTT CTATTATTCAGTTTCTATGTAC; tarM_d, GTCAGTCGACCTATTTGACTATTATCAACTTTCTTCGCTTTATG). The knockout of the *tarM2* gene caused the *S. epidermidis* strain E73 with two types of WTA to be sugar-less. A pRB474-*tarM2* plasmid was constructed using the pRB474 plasmid and two primers (tarM2_E73_F, GTCGGATCCAAAGGAGGTTATATAATGAAACAAACTTATATGATTGTA AATGAGTTGG, and tarM2_E73_R, CCGATGAATTCTTAGTTAAACAATTGTTTCCATTTCACCATC). This *tarM2* knockout strain was complemented with the pRB474-*tarM2* plasmid expressing α-glucose on poly(RboP)-WTA again (Appendix Fig 2a.). This *S. epidermidis*-originated *tarM2* gene could also be expressed in *S. aureus* strains. The expression of this novel *tarM2* gene in the no-sugar WTA *S. aureus* strain RN4220Δ*tarM*Δ*tarS* also resulted in the expression of a poly(RboP)-WTA with α-glucose (Appendix Fig 2b.).

a



b

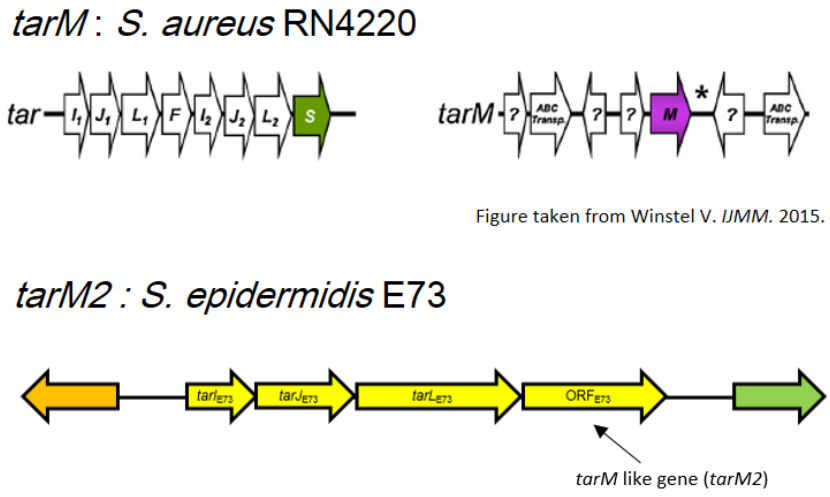
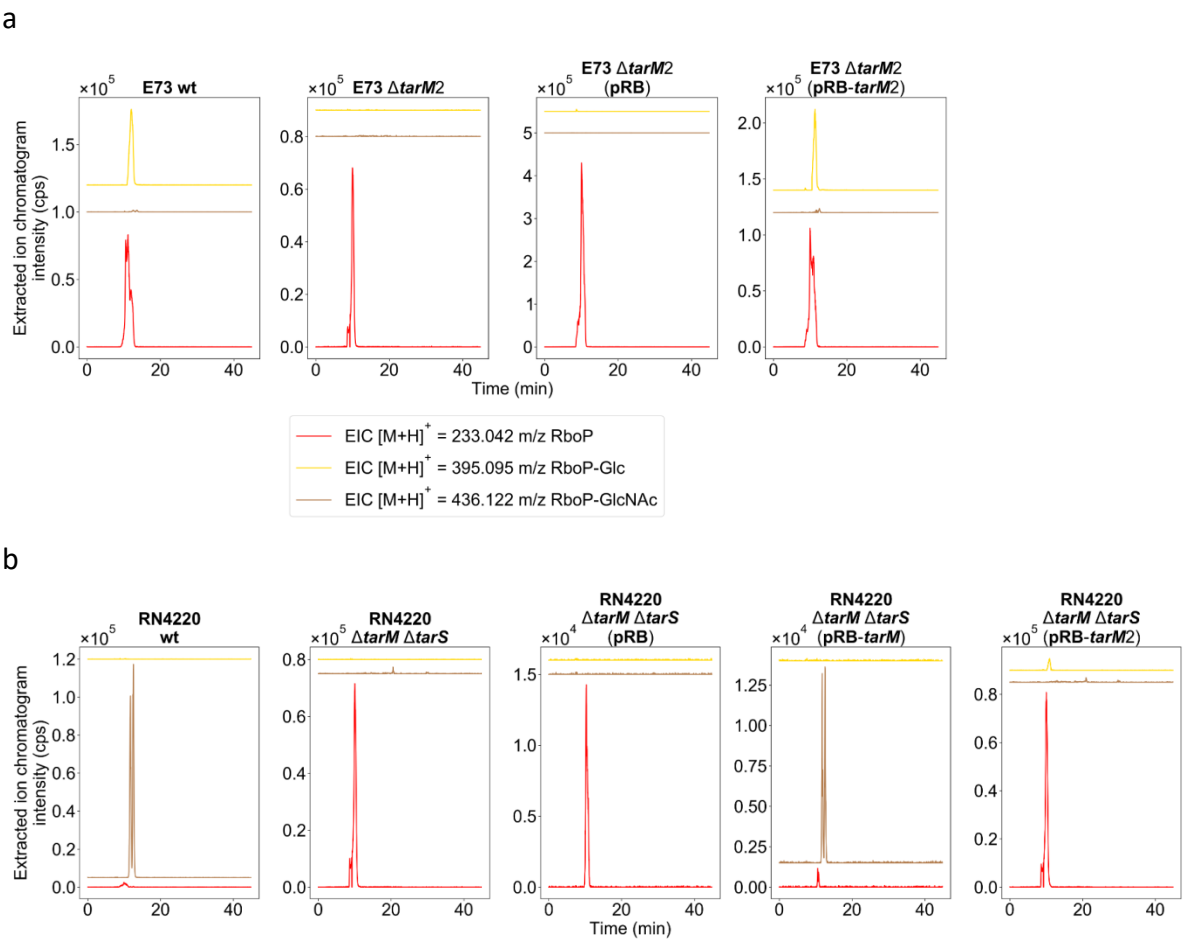


Figure taken from Winstel V. *IJMM.* 2015.

Appendix Figure 1. (a) Sequence alignment of TarM from *S. aureus* and the novel TarM2 from *S. epidermidis* showing a high amino acid sequence identity. The shaded (black) boxes marked the different amino acids of the two proteins. (b) Different locations in the genome of *tarM* and *tarM2*.

While the *tarS* gene is located just near the WTA polymer synthesis enzyme gene cluster *tarIJL*,

the *tarM* gene in *S. aureus* sits independently in another position far away from other WTA-related genes in the genome. The genome sequence of the strain E73 with two types of WTA revealed that the novel *tarM*-like gene *tarM2* was similar to the *tarS* gene in *S. aureus,* located just beside the *tarIJL* genes.



Appendix Figure 2. Liquid chromatography-mass spectrometry (LC-MS) showed that the *tarM2* gene plays an important role in the transfer of glucose to WTA. Counts per second (CPS) measured by LC-MS are shown. EIC: Extracted ion chromatogram intensity. (a) *S. epidermidis* E73 panel; (b) *S. aureus* RN4220 panel.

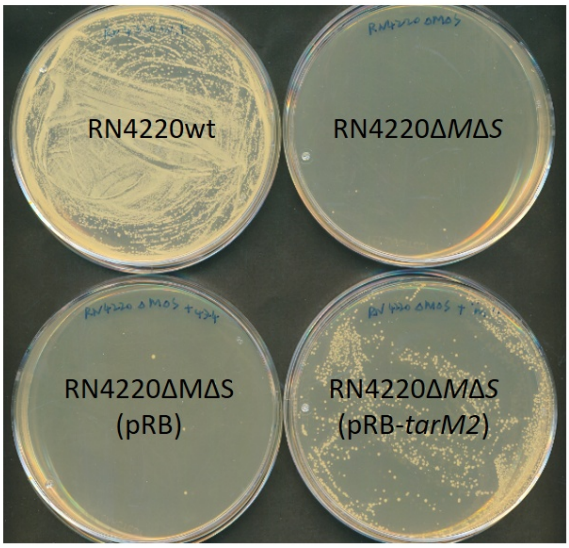
To confirm that α-glucose has a similar function in the interaction with phages, the phage spot assay and SaPI Φ11 assay were performed (Appendix Fig 3). The *spa* mutants were also tested because we wanted to test the same strains in the human IgG binding assay later. ΦK formed spots on all the strains because it could bind to the backbone. Φ11 formed spots only on the *tarM, tarM2*, or *tarS* complement strains of the no-sugar WTA *S. aureus* strain RN4220Δ*tarM*Δ*tarS*. The podovirus Φ68 formed spots only on the *tarS*-expressing strains without

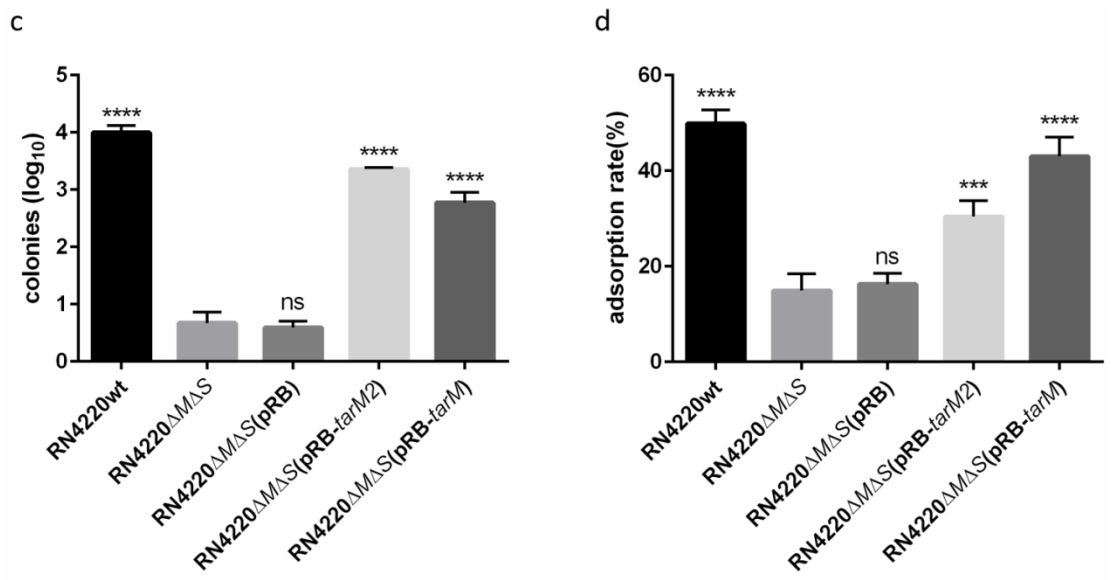
tarM or *tarM2* but not the strains with *tarM* or *tarM2*. This was the same result as in Chapter 4, showing that *S. aureus* could protect itself from the lysis of podoviruses by acquiring the *tarM* gene. Not all *S. aureus* strains have this *tarM* gene. These results indicate that the novel *S. epidermidis* TarM2 behaves similarly to the *S. aureus* TarM. *S. epidermidis* might be the origin *S. aureus* obtained the *tarM* gene from some time ago. The phage Φ11 could not discriminate α-glucose from α-GlcNAc.

a

Strains	ΦK	Φ11	Φ68	Φ187	ΦE72
RN4220Δ*spa*					
RN4220Δ*spa*Δ*MΔS*					
RN4220Δ*spa*Δ*MΔS* (pRB)					
RN4220Δ*spa*Δ*MΔS* (pRB-*tarM*)					
RN4220Δ*spa*Δ*MΔS* (pRB-*tarM2*)					
RN4220Δ*spa*Δ*MΔS* (pRB-*tarS*)					

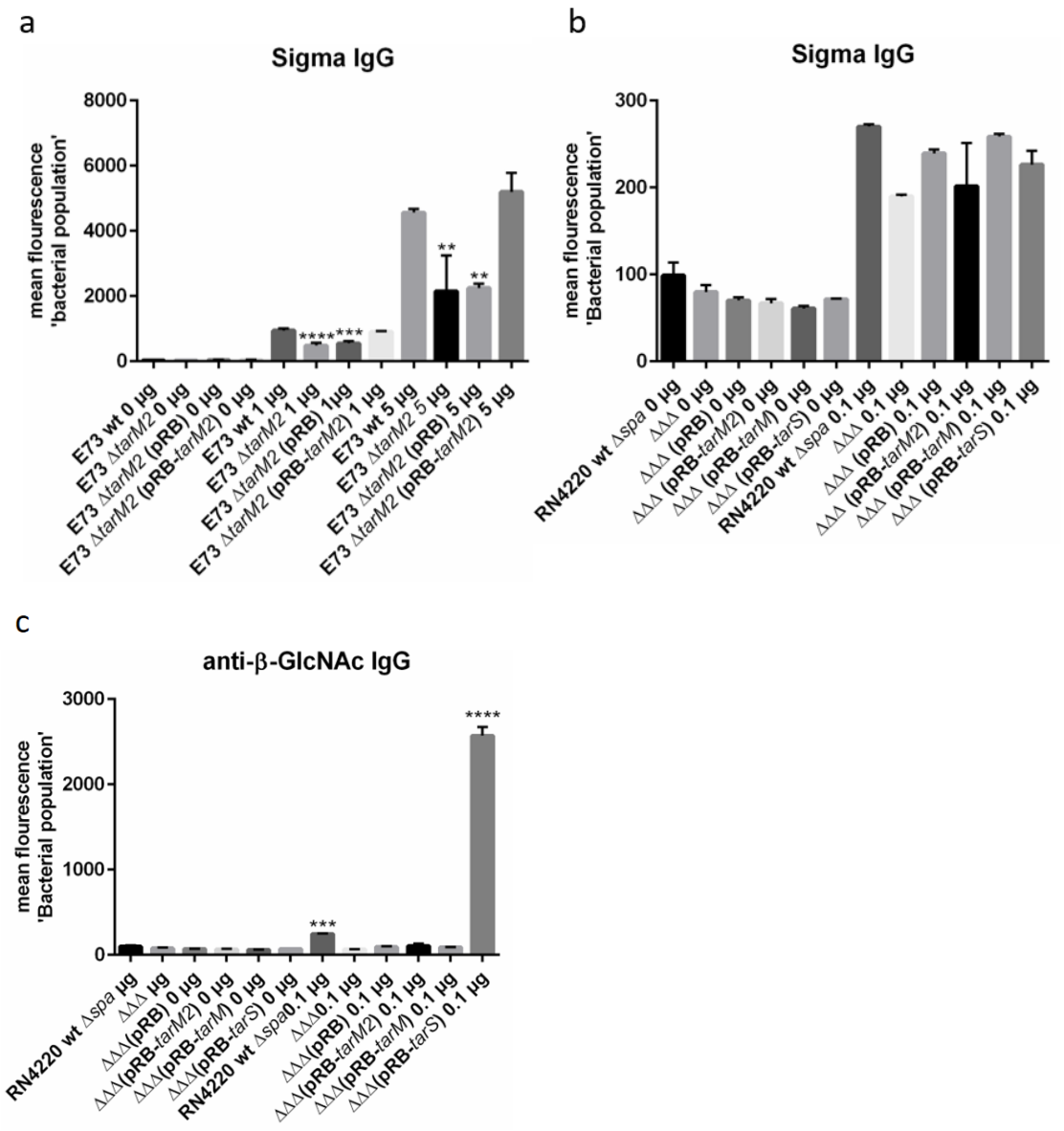
b





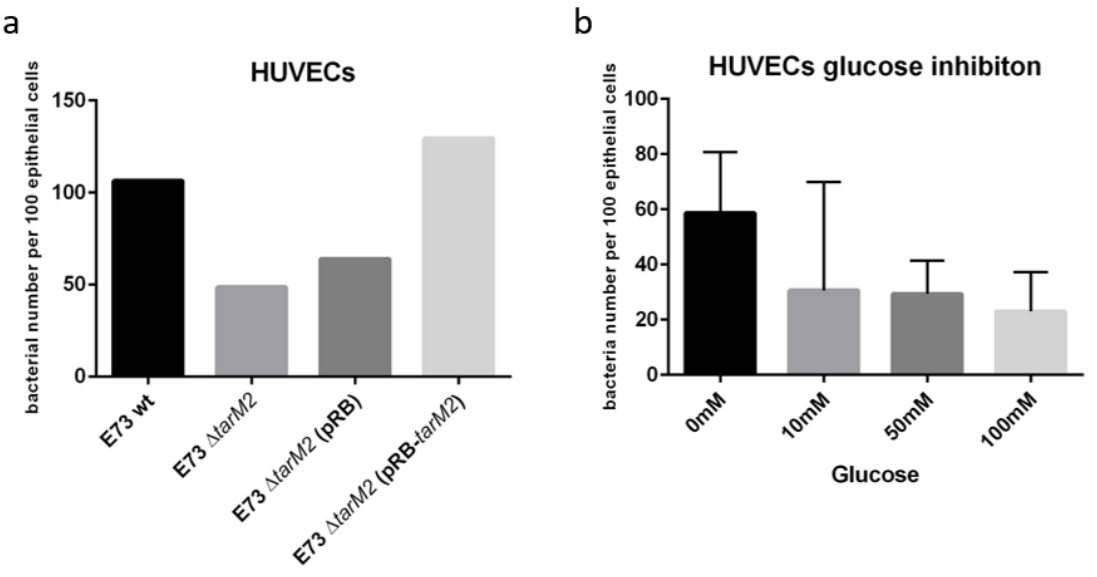
Appendix Figure 3. The *tarM2* gene caused Φ11 susceptibility of the bacteria. (a) Spot assay of different phages. The *tarM2* complement made the *S. aureus* RN4220 no-sugar mutant sensitive to Φ11 infection. (b) Bacterial colonies on TSA plates with 3 μg/ml tetracycline in the SaPI Φ11 assay. (c) Results of the SaPI Φ11 assay. (d) Results of the Φ11 adsorption assay. Statistical significance, when compared to Δ*tarM*Δ*tarS*, was analysed by one-way ANOVA with Dunnett's post-test. ns, *P*>0.05; *, *P*≤0.05; **, *P*≤0.01; ***, *P*≤0.001; ****, *P*≤0.0001.

To investigate whether *tarM2* could affect the recognition of bacteria by IgG in human serum, an *in vitro* IgG binding assay was performed. For all flow cytometry experiments, the bacterial panel lacking the *spa* gene and encoding the IgG-binding protein A was used. The results showed that *tarM2* functioned in *S. epidermidis* similar to *tarS* in *S. aureus,* which is highly immunogenic. The expression of *tarM2* led to significantly increased IgG binding compared to the glucosylation-deficient mutant E73d*M2*, indicating that glucose on the poly(RboP) backbone of WTA in *S. epidermidis* is an important epitope for human serum antibodies (Appendix Fig 4a). The *S. aureus* RN4220 panel was also tested with human serum in the IgG binding assay (Appendix Fig 4b). Although the trend was not significant, it was similar to the trend with the E73 panel. Specific anti-β-GlcNAc WTA IgG, a kind gift from Dr. Robin von Dalen, was also tested to determine if there was a cross-reaction of this glucosylated WTA with β-glycosylated WTA in IgG binding in the human immune system (Appendix Fig 4c). However, the results showed that this specific IgG could not recognize the glucosylated WTA even when expressed in *S. aureus*.

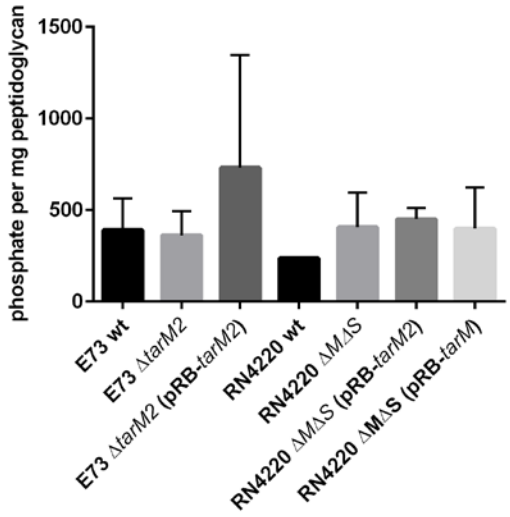


Appendix Figure 4. The knockout of the *tarM2* gene attenuates the immunogenicity of WTA. (a) Less IgG binds to the glucose mutant of *S. epidermidis* strain E73 with two types of WTA compared to the wild-type strain. (b) RN4220 panel detected using β-GlcNAc-specific IgG (n=2). (c) RN4220 panel detected using β-GlcNAc-specific IgG.

A previous study of the whole gene cluster *tarIJLM2* suggested that *tarIJLM2* could promote the invasiveness of *S. epidermidis* by increasing binding to human endothelial cells (see Chapter 2). The *tarM2* gene might also play a role in affecting bacterial binding ability. Thus, binding assays to different cells were performed (Appendix Fig 5).



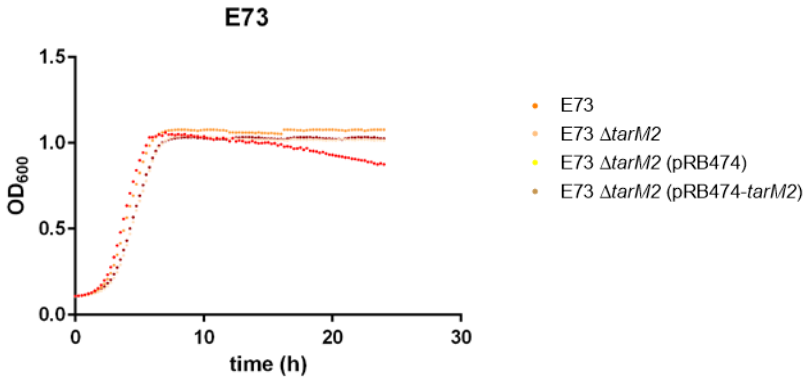
Appendix Figure 5. The *tarM2* gene might be an important factor in *S. epidermidis* binding to human endothelial cells. The glucose on RboP-WTA might play an important role in the binding process. (a) *In vitro* binding assay to HUVECs (n=1). (b) Additional preincubation of HUVECs with glucose decreased the binding of *S. epidermidis* with glucose-RboP WTA to the cells (n=2).



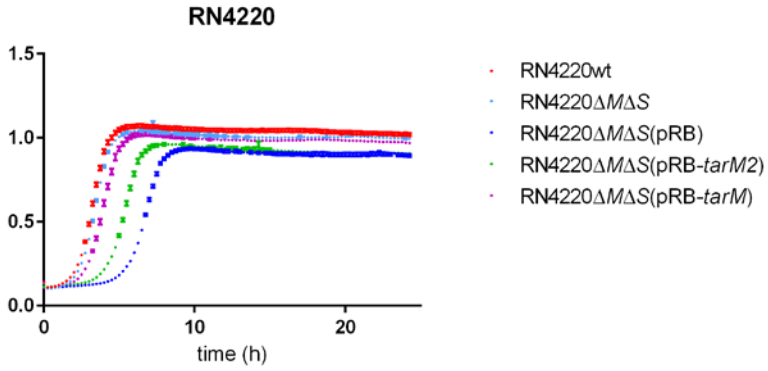
Appendix Figure 6. The knockout and complement of the *tarM2* gene in *S. epidermidis* E73 and *S. aureus* RN4220 did not affect the total amount of WTA on the surface of bacteria.

The growth curve showed that the knockout and complement of the *tarM2* gene did not influence the total amount of WTA per cell wall or the growth of bacteria (Appendix Fig 6, 7). The WTA of the wild-type strain and *tarM2* mutants was isolated and detected in a PAGE gel. However, there was no expected difference between the RN4220wt and *tarM2* complement of the no-sugar RN4220d*M*d*S* mutant (Appendix Fig 8). The effects of *tarM2* on biofilm formation and oxacillin resistance were also tested (results not shown). However, there were negative results in these two assays. This indicates that *tarM2* did not play a role in biofilm formation or oxacillin resistance. The expression plasmid pBAD-*tarM2* was also constructed with two primers (tarM2-pBAD up, ATGAAACAAACTTATATGATTGTAAATGAGTTG, and tarM2-pBAD dn, GTTAAACAATTGTTT CCATTTCACCATC), expressed in *E. coli* TOP10 and sent with all information to Dr. Yinglan Guo in the lab of Prof. Tillo Stehle for protein crystallization for structural analysis.

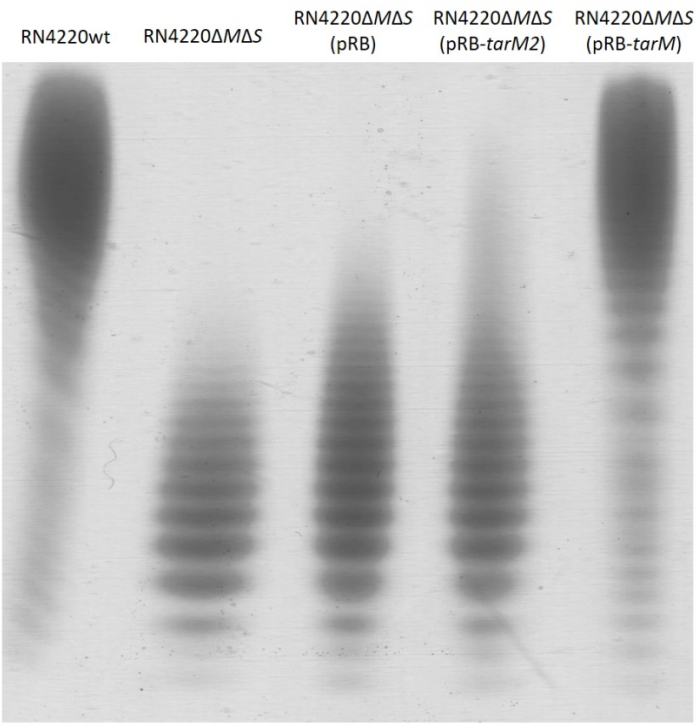
a



b

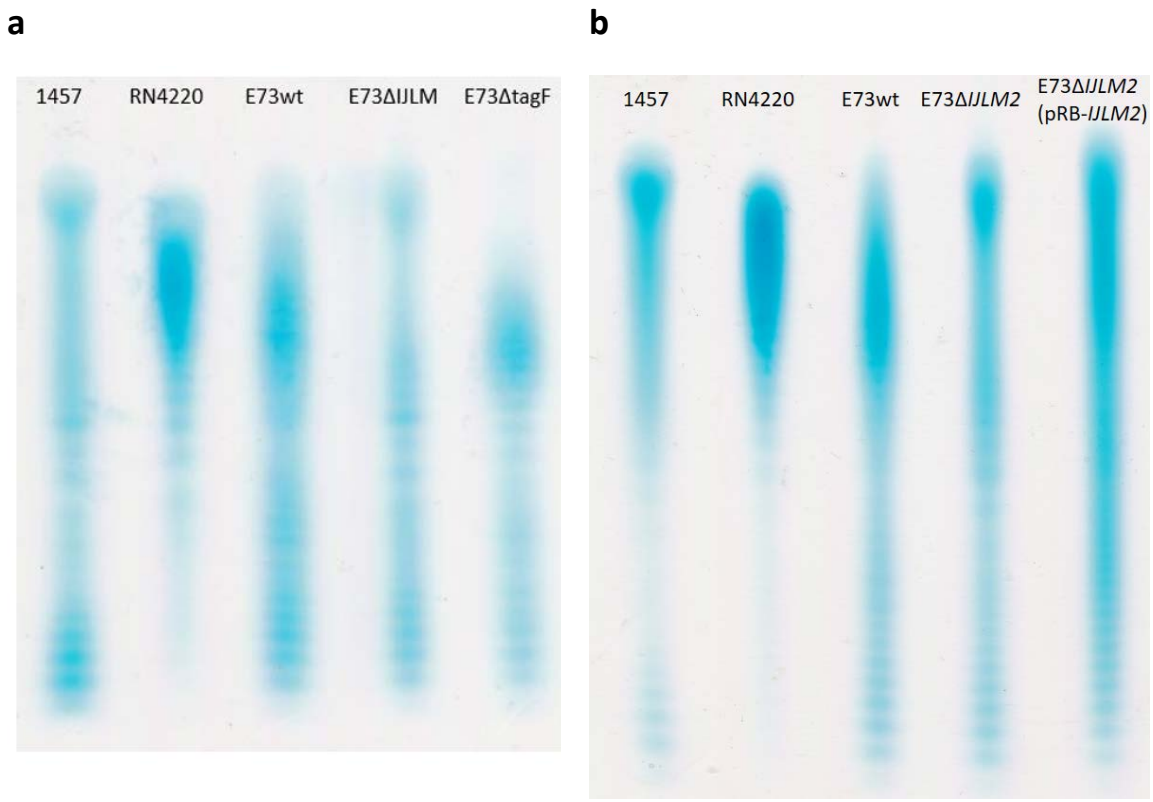


Appendix Figure 7. The knockout and complement of *tarM2* in *S. epidermidis* E73 (a) and *S. aureus* RN4220 (b) did not affect the growth of bacteria.

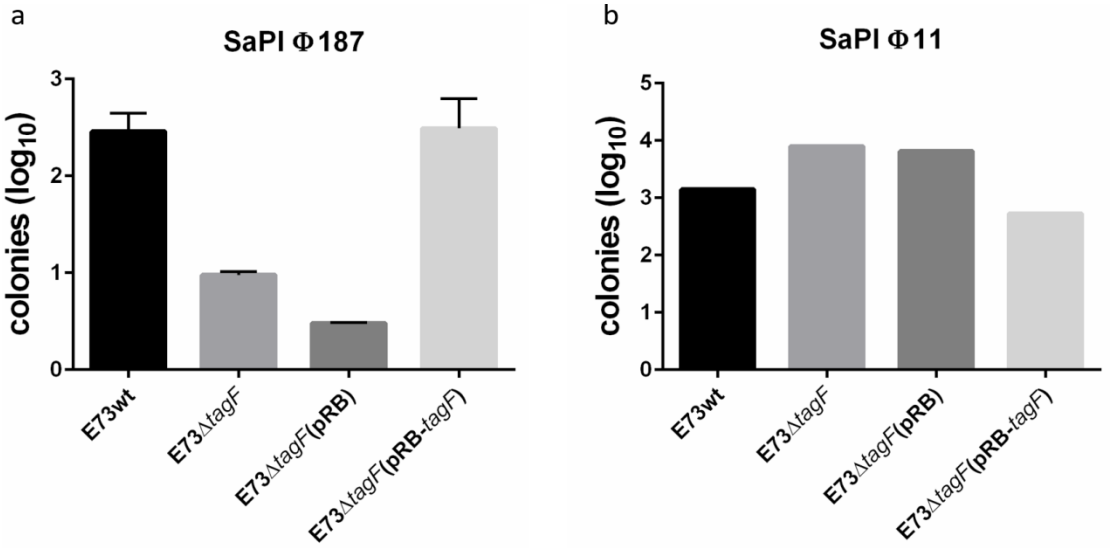


Appendix Figure 8. WTA-PAGE gel showed no difference in the WTA after complementation with the *tarM2* gene of the no-sugar mutant Δ*tarM*Δ*tarS* in *S. aureus* RN4220.

For the *S. epidermidis* strain E73 with two types of WTA. WTA of the wild-type strain and the mutant strains were isolated and detected in a PAGE gel (Appendix Fig 9). The WTA samples from *S. aureus* RN4220 and *S. epidermidis* 1457 were used as glycosylated RboP-WTA and glycosylated GroP-WTA controls, respectively. E73wt had both RboP-WTA with glucose and non-glycosylated GroP-WTA, which resulted in an emerged pattern of RN4220 and 1457. The *tarIJLM2* mutant of E73 was more similar to the WTA sample of *S. epidermidis* 1457 in the gel because its RboP-WTA was knocked out (Appendix Fig 9a). The E73wt strain was observed to have RboP-WTA with relatively smaller molecules when compared to the RboP-WTA from RN4220. The complementation of the *tarIJLM2* gene to the *tarIJLM2* mutant presented more WTA in the upper part, such as RN4220 RboP-WTA, but was not exactly complementary to the same pattern as E73wt (Appendix Fig 9b).



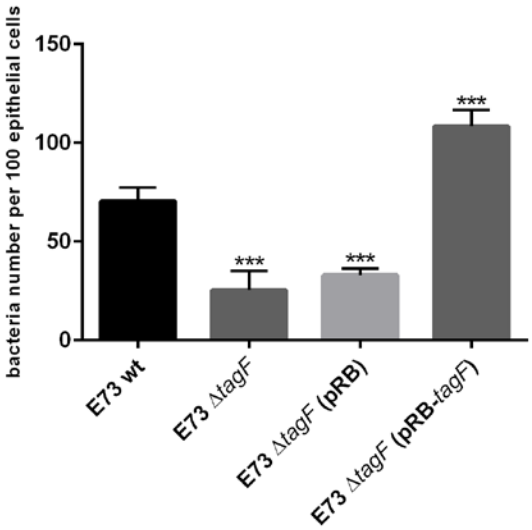
Appendix Figure 9. Two separate repeats of WTA-PAGE gel showed WTA samples with the same phosphate amount from different strains.



Appendix Figure. 10. SaPI assay with Φ187 (a) (n=2) and Φ11 (b) (n=1).

Results and discussion for the *tagF* gene in *S. epidermidis* with two types of WTA

To knock out the GroP-WTA from E73, the pBASE6 plasmid and four primers (tagF_a, CGATGGTACCCTACACTCCTCTAATGATGAATTGAATCA; tagF_b, CGTCCTTTCTCTTTATATTAAATCGACA AC; tagF_c, TAATATAAAGAGAAAGGACGGGGTATAATATTAATGAGCAATTAATATTATGCC; tagF_d, GTCA GTCGACCTTCATCTTATTGATGATTTTCAAAATAAAAGCG) were used. For the *tagF* complement strain, the pRB474 plasmid and two primers (tagF_up, GTCGGATCCAAAGGAGGTTATATA ATGAATAAACTGACAATTATTGTCACGTATTAT, and tagF_dn, CCGATGAATTCTCATTGTTCCTTGATATCC TTATGAATTAAATC) were used. Some GroP-WTA appeared to remain in the GroP-WTA knockout strain E73Δ*tagF* in the lower part of *S. epidermidis* 1457. Only the upper part of the GroP-WTA with larger molecules was missing after knockout of the *tagF* gene (Appendix Fig 9a). Therefore, the hypothesis could be that smaller GroP-WTA polymers with shorter chains remained. It could be that *tagB* or another second GroP polymerase is active after the knockout of the *tagF* gene in *S. epidermidis*. Additional evidence regarding the WTA structure was found by LC-MS and NMR analysis; the GroP-WTA could still be detected in E73Δ*tagF* (results not shown). However, the strain E73Δ*tagF* behaved similar to the *tagF* gene knockout in the SaPI assay (Appendix Fig 10). Moreover, when performing the cell binding assay using the *tagF* gene knockout mutant and complement, a significant reduction in bacteria binding to cotton rat nasal cells was observed in the GroP-WTA knockout strain E73Δ*tagF* when compared to the E73wt with two types of WTA (Appendix Fig 11). This indicated that the WTA structure of E73 in the *S. epidermidis* strain with two types of WTA changed after the knockout of the *tagF* gene, even though some GroP-WTA remained. However, from the LC-MS, we could roughly see that the amount of *tagF* did not changed. Therefore, a hypothesis was that the second unknown GroP-WTA polymerase makes shorter but more GroP-WTA when compared to the longer GroP-WTA from TagF, while it takes up more linkage units on the bacterial cell surface, decreasing the relative amount of RboP-WTA. Further research needs to be done to answer this question.



Appendix Figure 11. Human lung cell line A549 binding assay. Statistical significance when compared to wild type was analysed by one-way ANOVA with Dunnett's post-test. ns: *P*>0.05, *: *P*≤0.05, **: *P*≤0.01, ***: *P*≤0.001, ****: *P*≤0.0001.

Appendix-Method

WTA PAGE:

PAGE for WTA was performed as described previously with some modifications (Corzo *et al*., 1991; Meredith *et al*., 2008).

Preparing a 20 cm Gel:

-Resolving gel: mix 20 ml of 2 M Tris-HCl buffer (pH 8.2) with 40 ml of acrylamide stock solution (Rotiphorese gel 40, 19:1); use 600 µl of 10% ammonium persulfate and 60 µl of TEMED for polymerization.

-Stacking gel: mix 6 ml of ddH_2O with 1 ml of acrylamide stock solution and 3 ml of 2 M Tris-HCl (pH 8.5); for polymerization, use 100 µl of 10% ammonium persulfate and 10 µl of TEMED.

Run the WTA PAGE:

- Quantify the phosphate amount of the WTA samples with the phosphate assay before loading
- Load the WTA samples (approximately 100-200 nmol of Pi in 7.5 µl of water + 2.5 µl of 4 x Tris/Tricine loading
dye) into each lane
- For GroP-WTA use 300 V and 15 mA for 18 hours; for RboP-WTA, 300 volts and 40 mA for 18 h

Stain the gel:

- Fix the gel in EAW solution (40% EtOH; 5% HAc in ddH_2O) for 1 hour
- Stain the gel with the Alcian Blue solution (0.0375 g of Alcian Blue, 75 ml of 2% HAc, 75 of ml EAW) for more than 4 h
- Wash the gel twice with EAW solution for 1 h for the first time and overnight for the second time.

IgG binding assay:

- Exponentially growing bacterial cultures were adjusted to an OD600 of 0.5 and then diluted to 1:10 in PBS
- Then, 100 µl of diluted bacteria was mixed with 100 µl of IgG diluted in PBS with 1% BSA (10 µg/ml for IgG from pooled human serum purchased from Sigma). A control without IgG was included in all experiments for all mutants.
- Mixed samples were incubated at 4°C for 1 h, centrifuged, washed 2–3 times with PBS, and further incubated with 100 µl of FITC-labelled anti-human IgG at 4°C for 1 h.
- Bacteria were centrifuged, washed 3 times with PBS, and fixed with 2% paraformaldehyde (PFA)
- The IgG bound to the bacterial surface was quantified by flow cytometry.

Acknowledgement

I would like to express my deepest sense of gratitude to my tutor Prof. Andreas Peschel for his trust, encouragement and patient guidance leading me through the scientific journey during my Ph.D. project. Now I feel confident to do biological research at a more basic level to determine the mechanism of a clinical phenomenon.

To Dr. Volker Winstel, for his idea and continuous help. He is not only a teacher who has taught me most of the methods and knowledge I need for my Ph.D. research but also a great young scientist model who has demonstrated how to think comprehensively and keep the enthusiasm for research.

To the cell wall team, for the helping and friendly atmosphere. I would like to give the most thanks to our technician Regine Stemmler. She has truly helped me a lot with the lab work, especially during my pregnancy, even though she is quite busy with strain collection and delivery for the whole lab. She is always with a lovely smile that brings sunshine into our office. Janna Hauser is a happy girl always caring for others and me. Dr. David Gerlach and Dr. Robin von Dalen have helped me a lot with computer work and discussions on WTA questions. I also thank Petra Kühner and Dr. Xuehua Li for their kind help both inside and outside the lab. Thanks also to my students for the nice cooperation time with the lab work: Lena Mühlenbruch, Luqing Li, Marcel Beha, Minyue Qi, and Alexandra Fux.

To Prof. Friedrich Götz, Prof. Christiane Wolz, Dr. Bernhard Krismer, Dr. Doro Kretschmer, Dr. Ralf Rosenstein, and Dr. Simon Heilbronner, for their kindness in my research work and for helpful discussions, allowing me to learn a lot from them.

To Gisela Bauer-Haffter and Daniela Binder, for their help in the administration work. Since I am a non-German speaker, they always paid more attention to me.

To lab mates, now and in the old times, Angelika Jochim, Darya Belikova, Benjamin OrlandoTorres Salazar, Dr. Christoph Slavetinsky, Dr. Ana Jorge, Katja Schlatterer, Cordula Gekeler, Christian Beck, Dr. Alexander Zipperer, Dr. Dennis Hanzelmann, Dr. Dominik Bloes, Melanie Knobloch, Claudia Laux, and many others, for the happy and lovely time inside and outside the lab.

To my collaborators from other labs, Dr. Jesper Larsen, Prof. Christoph Mayer, Axel Walter, Prof. Holger Rohde, Dr. Anna Both, Prof. Min Li, Dr. Jessica Schade, Esther Lehmann, Dr. Christopher Weidenmair, Dr. Patricia Sanchez Carballo, Dr. Ivana Maslanova, Dr. Matthias Marschal, Dr. Filipp Oesterhelt, and many others, for the great cooperative work in the past years. I also learnt a lot from them all.

Finally, I would like to give my special thanks to my family, my parents and my love Ye Tian, for their support and encouragement during my Ph.D. study. The little angel in my belly is also fulfilling her mother with love and power to create a better life in the future. I love you all!