Natural phage communities crosslink different species within the genus Staphylococcus

Pauline Goeller
ETHZ
Tabea Elsener
ETHZ
Dominique Lorgé
Swiss Federal Institute of Technology in Zurich
Natasa Radulovic
ETHZ
Viona Bernardi
ETHZ
Annika Naumann
ETHZ
Nesrine Amri
ETHZ
Ekaterina Khatchatourova
ETHZ
Felipe Hernandes Coutinho
Universidad Miguel Hernández
Martin Loessner
ETH Zurich https://orcid.org/0000-0002-8162-2631
Elena Gómez-Sanz (elena.gomez@hest.ethz.ch)
ETHZ https://orcid.org/0000-0001-7666-3642

Article

Keywords: Bacteriophage Host Range, Bacterial Communities, Genome Diversity, Wastewater, Genetic Mobilization, Host Specialism, Horizontal Gene Transfer

Posted Date: April 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-388089/v1

Version of Record: A version of this preprint was published at Nature Communications on November 29th, 2021. See the published version at https://doi.org/10.1038/s41467-021-27037-6.

Natural phage communities crosslink different species within the genus *Staphylococcus*

Pauline C. Göller[1], Tabea Elsener, Dominic Lorgé, Natasa Radulovic, Viona Bernardi, Annika Naumann, Nesrine Amri, Ekaterina Khatchatourova, Felipe Hernandes Coutinho[3], Martin J. Loessner[1], Elena Gómez-Sanz[1,2]*

[1]Institute of Food, Nutrition and Health, ETH Zurich, 8092 Zurich, Switzerland.

[2]Área de Microbiología Molecular, Centro de Investigación Biomédica de La Rioja (CIBIR), Logroño, Spain.

[3] Departamento de Producción Vegetal y Microbiología, Universidad Miguel Hernández, San Juan de Alicante, Spain.

*Correspondance: elena.gomez@hest.ethz.ch

Abstract

The importance of the bacteriophage host range builds on its role as an innate barrier, which defines the phages' impact on bacterial communities and genome diversity. Yet, little is known about host range natural patterns. We characterize 94 novel staphylococcal phages from wastewater and establish their host range on a diversified panel of 117 staphylococci from 29 species. Using this high-resolution phage-bacteria interaction matrix, we unveil a multi-species host range as a dominant trait of the isolated staphylococcal phages. Phage genome sequencing shows this pattern to prevail irrespective of taxonomy. Network analysis between phage-infected bacteria revealed that hosts from multiple species, ecosystems, and drug-resistance phenotypes share numerous phages. This could promote genetic mobilization facilitated by many transfer routes. Lastly, we demonstrate that phages throughout this network package foreign genetic material at various frequencies. Our findings defy a strong host specialism of phages and highlight great possibilities for horizontal gene transfer.

Introduction

Bacteriophages (phages) are the most abundant biological entities on Earth, and yet the understanding of the association between bacteria and their infecting phages remains limited. Host range interactions have profound implications on how phages influence bacterial community composition and ecology[1,2], or facilitate horizontal gene transfer[3-5]. Thus, the host range is a central trait to understand in phage biology (reviewed in[6]), which knowledge has important applications in industry and human health[7]. The global health crisis of drug-resistant pathogens could be ameliorated by phage therapy, a promising treatment strategy especially for antibiotic resistant bacteria[8]. Paradoxically, phages may impose adverse implications, as they could be suitable vectors for bacterial adaptation traits, such as virulence and antimicrobial resistance determinants[9]. However, their exact role in the exchange of genetic material remains unclear as transduction frequencies and the phages' range of influence is still unsettled. Per definition, the phage host range refers to the taxonomic breadth of bacteria a phage can successfully infect (reviewed in[10]). Labor-intensive infection assays showed that host ranges diversify from narrow to broad[11]. While "broad" and "narrow" are partially conditioned by the genetic diversity of challenged hosts, a narrow host range is commonly reported for phages that replicate on few hosts. In contrast, broad host range phages complete their lifecycle in numerous strains from distinct or the same species[6]. To date, most isolated phages (> 85 %) belong to the order *Caudovirales*[12] and are reported as specialists with narrow host ranges. On average, they infect two strains from a single species[2]. However, a few network studies from a compilation of published data or marine viruses find that phages can infect a multitude of hosts and that different phage types predate each bacterial species[13-15]. These data challenge a strong phage-host specialization and the reported frequency of narrow host range phages.

Bacteria of the genus *Staphylococcus* are part of the natural skin microbiota of mammals and life-threatening pathogens due to their increasing virulence and antibiotic resistance. Therefore, they are considered as targets for a phage therapy approach. Adversely, phage transduction, in particular generalized transduction, is increasingly perceived as the primary route for mobilizing antibiotic resistance within this genus (reviewed in[16]). Based on their ability to produce coagulase, staphylococci are divided into the traditionally more pathogenic coagulase-positive staphylococci (CoPS), with *S. aureus* as the major species, and coagulase-negative staphylococci (CoNS) such as *S. epidermidis*. Today, CoNS are recognized as major nosocomial pathogens with limited treatment options due to a large proportion of antibiotic resistant strains[17]. They are regarded as important reservoirs of antimicrobial resistance determinants that could spread to clinical most critical species. Dissociated from clinical manifestations and based on multilocus data, a refined phylogeny for staphylococcal species into 15 species cluster and six species groups was suggested recently[18].

Despite the significance of CoNS, most staphylococcal phage genomes deposited in the RefSeq database are phages isolated on *S. aureus*. In fact, of the 292 phages, only 31 phages derived from CoNS. Oliveira and colleagues showed recently that phages with siphoviral morphology dominate among the published staphylococcal phages. They were divided into two phylogenetic clusters; the highly represented cluster B with temperate phages of genome sizes ~40 kb, and cluster D, with a few presumably virulent siphoviruses of genome sizes ~90 kb. The virulent podo- and myoviruses of the genus *Staphylococcus* are distributed within phylogenetic clusters A and C, and feature genomes of either below 20 kb or greater than 120 kb, respectively[19].

The staphylococcal phage host range is thought to be defined by a hierarchical combination of host factors (reviewed in[20]). At the highest level, the availability of the phage receptor on the surface of the bacterium restricts binding and infection. Wall teichoic acids (WTAs) have been reported as primary targets[20]. They show high intra-species conservation[21] but diversify among different staphylococcal species[22,23]. This evolutionary divergence is postulated to restrict most phage infections and transduction across the species barrier[24]. Yet, a few reported staphylococcal phages infect diverse species[25-28], and a close relationship between CoNS and CoPS phages was described[29-31]. One level below, individual internal defense mechanisms such as restriction modification systems, CRISPR/Cas, or resident prophages narrow the host range so that a phage can infect some, but never all strains of a species[10,20]. Nevertheless, a well-defined picture of the staphylococcal phage host range is absent, especially when considering non-S. *aureus* species.

To better appreciate the natural host range of staphylococcal phages, we characterized 94 phages isolated from a wastewater treatment plant (WWTP). For each member of this natural phage community, we assessed the host range on a diverse set of 117 staphylococci originating from 29 species, and six presumably negative control strains from the *Macrococcus* and *Enterococcus* genera. This host array was selected to constitute a diverse community of human, veterinary, or environmental isolated bacteria with different clinical relevance and antimicrobial resistance profiles. We show that staphylococcal phages from various taxonomic groups, morphologies, and lifestyles infect hosts across species barriers, and unveil a phage pool able to incorporate foreign genetic material. We further demonstrate that strains from different staphylococcal species, ecosystems, or drug resistant phenotypes are closely connected through diverse phages. Our findings challenge the commonly reported host specialism of phages and place phages as potent vehicles for bacterial genetic exchange.

Results and Discussion

Host enrichment cocktails unveil a great abundance of staphylococcal phages in wastewater

To study the host range of phages, we isolated native staphylococcal phages from the influent and effluent of a WWTP in Zürich, Switzerland. Before phage isolation, we assembled enrichment cocktails of staphylococcal hosts that were selected to produce a diverse community. We assured the growth of each host within an enrichment cocktail by excluding bacteria harboring cross-infecting prophages or bacteriocin producers. Remarkably, out of the 76 staphylococcal strains tested, 25 interfered with bacterial multiplication of selected cocktail members through either inducible prophages (19) or bacteriocin production (6). Five additional strains were removed because of a low growth rate. Based on the remaining 46 strains, we established five enrichment cocktails (A-D), each consisting of eight to eleven staphylococcal hosts from different species (17), origins (19 environmental, 14 animal, three human, and ten unknown), and coagulase groups (16 CoPS and 30 CoNS) (Supplementary Table 1). Using those cocktails, we isolated 155 phages, 134 from the wastewater influent and 21 from the effluent, with varying efficiency (Supplementary Table 2). Overall, the enrichment hosts spanned types from phage permissive to phage resistant. We isolated phages on 26 phage-permissive strains from 15 staphylococcal species (Supplementary Table 3). Of the initial 30 CoNS enrichment hosts, 22 strains were phage susceptible, in contrast to only four of the 16 CoPS hosts. This outline resulted in the isolation of 136 phages on CoNS, but only a few on CoPS (19 phages) (Supplementary Table 4). Another 24 presumably temperate phages were isolated from wastewater bacterial lysogens by induction. The detection and thus isolation hosts for those phages were mostly *S. epidermidis*, and one *S. sciuri* strain (Supplementary Table 5). Interestingly, they also proved highly successful in segregating free phage particles when used in enrichment cocktails (Supplementary Table 3), making them good candidates for further isolation advances.

To sum up, we demonstrate the prevalence of staphylococcal phages in wastewater through the isolation of 179 staphylococcal phages on 15 different staphylococcal species (Supplementary Table 6, Column A-E). As the vast majority of recovered phages (160) were isolated on CoNS (23 strains from 12 species), and the minority (19) on CoPS (4 strains from 3 species) (Supplementary Table 7, Column A-C), we substantially increase the phage landscape of the genus *Staphylococcus*.

Most phages have a single isolation host

To unravel the host ranges of the isolated staphylococcal phages, we selected a diverse host panel of 123 bacteria from 32 different species, including 29 *Staphylococcus* (117), two *Macrococcus* (4), and one *Enterococcus* (2) species. The chosen hosts originated from either

the human (40), veterinary (53), or environmental (23) biome and presented a multi-drug resistant (35), resistant (48), or antibiotic susceptible phenotype (40) (Supplementary Table 8). First, the host range served to discriminate between isolated phages, as we considered phages with an equal host range on this array as identical. This characterization resulted in the collapse of the 179 isolated into 94 unique phages (Supplementary Table 6, column F). Of these, 80 phages were recovered from the influent and eight from the effluent. Effluent phages that were also isolated from the influent (4) were assigned to the outlet phage fraction. Six different phages further remained from the induction of bacterial lysogens. As we observed an almost 50 % redundancy in phage isolation, we analyzed their recovery frequency. To our surprise, 62 out of the 94 unique phages were isolated only once (1 plaque), whereas 32 were isolated between two and 16 times. However, this re-isolation did not arise from an excess of isolation hosts, as 76 phages had only one matching isolation strain, and 84 phages were recovered on only one species (Supplementary Table 9). Furthermore, the inclusion of multiple enrichment cocktails was beneficial, as > 80 % of the enriched phages (77) were isolated from only a single cocktail (Supplementary Table 10). Altogether, these findings highlight the importance of using a comprehensive panel of bacterial hosts for phage enrichment.

The phage-bacteria incidence matrix is intermediately modular and nested

Next, we analyzed the individual infection patterns of the 94 distinguished phages (Supplementary Table 8). These data gave rise to a large phage-bacteria interaction matrix with 1,135 positive infection outcomes of possible 11,562 interactions (Figure 1, Supplementary Table 11). We measured high-order properties of this phage-host biadjacency matrix, specifically modularity and nestedness. A nested network structure is evoked if phage host ranges build subsets of each other. The most specialized phage infects only the most permissive bacteria, and broader host range phages evolve to infect less permissive hosts without losing the ability to replicate on the ancestor. A recent re-evaluation of phage-host interaction matrices found that phage-bacteria networks are typically nested[15]. Modularity is a characteristic of phage-bacteria infection networks where groups of phages specialize on non-overlapping groups of hosts. It is associated with taxonomy and elicited when a large taxonomic diversity of bacteria is challenged[14]. We expected many diverse modules in our interaction matrix, as we impose bacterial hosts of a large species variety and geographic scale.

Supplementary Figure 1 shows the modularity (a) and nestedness (b) sorting of our sample matrix. We observe an intermediate situation in which neither clear modules nor a strictly nested condition emerged. The calculated nestedness (NODF = 40.91) is significantly higher than expected from a random matrix (z-value = 47.34, $\bar{x}_{distribution}$ = 21.15, 95% = 21.84, p=0.0099) but sorting has not resulted in a clearly nested structure. Similarly, the calculated

modularity Q = 0.38 is significantly higher than expected at random (z-value = 47.87, $\bar{x}_{distribution}$ = 0.21, 95% = 0.22, p=0.0099). However, 348/1135 = 30.6 % of the interaction occur between the four detected modules. Interestingly, each module consists of host strains from at least three different phylogenetic species groups[18] and four distinct staphylococcal species. Phage permissive hosts from individual species, however, were mostly limited to one module. Only strains from *S. sciuri*, *S. aureus*, *S. haemolyticus*, *S. xylosus*, and *S. lentus* were split between different modules (Supplementary Table 12). We conclude that on the genus level, strains of equal species tend to cluster within a module, whereby individual species do not build modules. Furthermore, the species composition within a module seems unrelated to their phylogenetic relationship. Overall, the observed pattern suggests a limited specialization of staphylococcal phages on individual staphylococcal species.

Broad host range is a prevailing trait of the isolated staphylococcal phages

Using the biadjacency matrix, we next sought to analyze phage predation. Overall, isolated phages infected only staphylococcal bacteria. We observed a remarkable high level of infectivity at the species level, as 27 of 29 staphylococcal species were infected. On a strain level, we find an almost equal number of phage-permissive and phage-resistant hosts, with 60 (51.28 %) of the 117 challenged staphylococcal strains tolerating phage infection. Consistent with the isolation of phages on mainly CoNS hosts, we observed a clear preference of infection on this bacterial group (89 % of all infections). Most challenged CoNS strains showed phage susceptible (49 of 68 strains), which stands in contrast to the CoPS (11 of 49 strains) (Supplementary Table 13, Figure 2). Environmental strains were most permissive (74 %), followed by animal (53 %) and lastly human (23 %) (Table 1, Supplementary Figure 2). Nevertheless, the three hosts with the highest phage predation were of animal origin: *S. lugdunensis* I0507 (CoNS), *S. schleiferi* I3823 (CoPS), and *S. epidermidis* I0564 (CoNS) which were permissive for 65, 63, and 58 different phages, respectively. Generally, strains and species were infected by multiple phages, as on average, they were susceptible to 9.2 ± 15.4 (n = 123) and 22.7 ± 22.2 (n = 32) different phages, respectively.

Traditionally, staphylococcal phages are reported as species specific with a narrow host range. Here, we unveil that phages infect 12.0 ± 5.4 (n = 94) strains from 7.7 ± 3.7 (n = 94) species on average. In fact, the host range of 90 phages in this natural community spans multiple species, and only four phages exclusively replicated on a single species. Among them, three phages (PG-2021_89, PG-2021_93, and PG-2021_94 on *S. epidermidis*) were isolated by induction and feature a temperate lifestyle (see genomic data below). Hence, their detected plaquing host range may not reflect the true underlying host range. The remaining species-specific phage, PG-2021_6, was isolated from the outlet fraction and plaqued on a single strain (*S. sciuri*). On the other end of the spectrum, PG-2021_17 displayed the broadest lytic

potential. This is the sole phage isolated on *S. pseudintermedius,* and infected 32 strains of both CoPS and CoNS from 18 different species. Generally, we find that the host range of most isolated phages (86 %) spanned CoPS and CoNS, of which 22 phages covered the two clinically relevant *S. aureus* and *S. epidermidis*, and on average, another 10 ± 2 (n = 22) different species. Furthermore, all 90 broad host range phages infected strains of at least two different staphylococcal species groups[18]. On average, each phage infected strains from 3.4 ± 0.9 (n = 94) different species groups and six phages replicated on five out of the six possible groups.

Our findings seem inconsistent with the commonly reported phage specialization[20,32]. However, we hypothesized that specialization does not necessarily contradict a broad multi-species host range, as polyvalent phages can predominantly infect strains of a single species. A prevalent example is phage K, which is reported as an *S. aureus* phage that replicates on a few other staphylococcal species[25]. On the selected host panel, phage K infected 29 strains from 12 different staphylococcal species. Nevertheless, *S. aureus* hosts (15, 51.7 %) predominantly composed its host range (Figure 3). We evaluated whether a similar proportion of infected strains among the here characterized broad host range phages prevailed. Surprisingly, only 30 phages revealed a species tendency, with ≥ 50 % of all infected hosts belonging to one individual species. Those phages favorably replicated on *S. epidermidis* (25) and *S. sciuri* (5) (Supplementary Table 14, Column E). On the contrary, the established host range for 60 (64 %) of our broad host range phages had no apparent centralization of infection (Figure 3, Supplementary Figure 3).

Our data challenge a strong species tropism of phages within the genus *Staphylococcus* and excludes a harsh species boundary for staphylococcal phages. However, one must consider that the taxonomic diversity of bacteria greatly influences species specificity in each host array. Thus, host range proportions might shift in different collections with an equal number of strains per species.

Staphylococcal phages infect antibiotic-resistant strains from different biomes

We further employed the established interaction matrix to examine whether staphylococcal phages replicate on antibiotic resistant strains isolated from the environment, clinic, or veterinary biome. In total, we find 65 % of the antibiotic-susceptible, 50 % of antibiotic-resistant, and 29 % of the multidrug-resistant strains permissive to the tested phages (Supplementary Figure 2). When combining antibiotic resistant and multidrug-resistant strains, forty-one percent were infected by at least one phage. In addition, almost half of all infections (44.4 %) in the interaction matrix pertained to this group of hosts. Thus, similar phage predation occurred between antimicrobial susceptible and resistant hosts (two-sided Wilcoxon rank sum test with continuity correction, W=568.5, n = 60, p-value=0.068; Supplementary Figure 5).

Ultimately, all isolated phages productively infected at least one drug-resistant strain. Our results evidence that infection of antibiotic-resistant strains by phages from anthropogenic environments is common.

Bacteriophages could be suitable vectors for genetic exchange due to their vast abundance, stability in the environment, and their ability to bridge the spatial separation of donor and recipient bacteria[9]. The infection of hosts from diverse ecosystems is thereby a pre-requisite. To assess the potential phage-induced transfer of genetic material, we tested the phages' ability to connect hosts from the environmental, veterinary, or human biome. Only two phages infected bacteria from a single isolation origin, whereas all other 92 phages infected strains from at least two ecosystems. Within those, the host range of 58 phages connected veterinary and environmental isolated staphylococci and another 34 phages integrated strains recovered from humans. Interestingly, those 34 also infected drug-resistant and susceptible bacteria (Supplementary Figure 6a) and displayed distinct phage morphologies and lifestyles; both virulent and temperate (Supplementary Table 14).

In conclusion, we show that diverse staphylococcal phages connect naturally occurring hosts from different ecosystems and drug resistance phenotypes, suggesting this feature to be a general competence.

Natural phage communities crosslink species within the genus *Staphylococcus*

Next, we analyzed the established biadjacency matrix focusing on the interplay between infected bacteria rather than individual phage host ranges. To do so, we reduced the matrix to the 60 phage-permissive strains, which represented 27 different staphylococcal species. The network was collapsed into a bipartite projection in which hosts are represented as nodes and phages as edges. An edge between two bacterial nodes indicates the presence of at least one phage infecting both hosts, and the edges are weighted according to the number of phages that do so. The projection showed an interconnected network with 1,030 host interactions through 93 different staphylococcal phages (connectance = 0.58) (Table 3, Figure 4). We sought to establish parameters that best describe how this natural phage community crosslinks members of the genus *Staphylococcus*. On the one hand, we consider the number of shared phages between two hosts as an important marker, as they indicate transfer routes and opportunities for genetic exchange. Thus, the higher the number of shared phages between two hosts, the higher the chance of genetic displacement as multiple phages could govern a transfer. On the other hand, we recognize the number of direct neighbors, which is the count of nodes connected by an edge to the specified node. Neighbors are a measure of centrality and demonstrate the host's impact.

The bipartite projection revealed that staphylococcal strains from different species groups[18] share on average 3.8 ± 7.3 (n = 1293) phages. Moreover, individual staphylococcal strains and

species were connected by 4.3 ± 7.9 (n = 1770) (Supplementary Figure 4) and 4.1 ± 7.5 (n = 1666) different phages, respectively. However, staphylococcal strains from the same species were significantly better connected (two-sided Wilcoxon rank sum test with continuity correction: W = 237492, p-value = 1.93 e -15), as on average they were linked by 8.3 ± 11.4 (n = 104) different phages. Surprisingly, the two best connected hosts throughout this network belong to different species, species groups[18], and coagulase types: *S. lugdunensis* I0507 (Epidermidis-Aureus, CoNS) and *S. schleiferi* I3823 (Hyicus-Intermedius, CoPS), which share sensitivity to 58 different staphylococcal phages.

In addition to the numerous transfer opportunities, we found a bacterium to be connected to 34.3 ± 13.6 (n = 60) strains from 17.4 ± 5.2 (n = 60) staphylococcal species through phages. The strain with the highest number of neighbours is most likely to receive and donate genetic material. We found *S. vitulinus* C5817 as the most central host that could interact with 56 of 59 available hosts. Furthermore, phage infections connected strains of the species *S. epidermidis* (I0564), *S. lugdunensis* (I0507), and *S. schleiferi* (I3823) to 25 of 26 other staphylococcal species.

Lastly, we appraised the connectivity between ecosystems by phages, as there is a rising fear of genetic mobilization between the human, environmental and animal biome. To do so, we assessed the number of shared phages between hosts of different origins. Surprisingly, we found no significant difference in the average number of shared phages between hosts from the same (4.7 ± 8.4, n = 565) or different biome (4.1 ± 7.6, n = 1205) (Two-sided Wilcoxon rank sum test with continuity correction, W=1315504, p = 0.08981). Hosts of environmental and veterinary origin, however, were exceptional well connected, as they share 5.7 ± 8.8 (n = 476) phages on average (Supplementary Figure 6b). Furthermore, of the on average 34 neighbors previously found for a host in this network, only 11.2 ± 7.0 (n = 60) share the same isolation biome, whereas 23.1 ± 9.9 (n = 60) hosts derived from different ecosystems (Supplementary Figure 7). The interconnection of spatially separated staphylococcal strains becomes critical when addressing the dissemination of drug resistance determinants. Here, we demonstrate that each drug resistant host is connected on average to 16.0 ± 6.7 (n = 33) drug susceptible neighbors through 4.3 ± 8.0 (n = 891) different phages. Our findings evidence the existence of multiple routes and opportunities for genetic material to be mobilized by phages between hosts of different species, sources, and clinical relevance.

WWTPs are reservoirs for diverse CoNS phages

We sequenced the genome of 40 CoNS viruses of our natural phage community (Table 2) and assessed their morphology by electron microscopy (Figure 5). Among the 40 sequenced phages, 29 were isolated from the WWTP inlet, seven from the outlet, and four were bacterial lysogens. Overall, we identified 29 myoviral and 11 siphoviral morphologies. Isolated phages from the raw wastewater revealed to be mainly myoviruses (with two siphoviruses), whereas

siphoviruses dominated in the treated water. All induced prophages were siphoviruses (Table 2). As anticipated, the sequenced myovirus' genome sizes ranged from 128.3 - 145.1 kb, while the siphoviruses separated into two groups between 42.2 - 44.5 kb and 85.8 - 92.2 kb[19]. Interestingly, all siphoviruses with a larger genome were isolated as free viral particles and displayed the distinct morphology with tails > 300 nm, while the smaller ones were solely isolated after induction (Figure 5). Lysogeny modules were only found in the genome of the latter. This is coherent with literature, as smaller staphylococcal siphoviruses are predicted to be temperate, whereas larger siphoviruses are presumably virulent[19]. To date, only three representatives of the latter are reported. With the characterization of seven novel large siphoviruses, we significantly extend the currently available sequencing landscape of this phage fraction. Next, we assigned the closest phage relative for each of our novel phages based on average nucleotide identity (ANI). Interestingly, 29 CoNS viruses shared a relatively high genome identity (> 88 % ANI) with known staphylococcal phages, while the other 11 appeared to be distantly related (< 70 % ANI). We detected a total of 34 tRNAs among 18 phage genomes. All tRNAs-encoding phages corresponded to strictly lytic myoviruses or large siphoviruses. These results are compatible with the hypothesis that tRNAs are more prevalent among virulent phages. They are less well adapted to their replication hosts and hence, have a compositional difference for codon or amino acid usage[33]. Lastly, we computed a phylogenomic analysis using the phage genomes described herein along with 292 staphylococcal phages deposited on NCBI (Figure 6). As a unique ecosystem, water from a WWTP revealed to contain diverse staphylococcal phages from different families and genera. The phylogenomic tree showed a good agreement between phage morphology, genome length, and taxonomy. However, the extent of the phage host range seemed rather independent, although members of the *Herelleviridae* infected the highest number of strains and species, followed by ~ 90kb, and lastly, ~40 kb *Siphoviridae* (Figure 6). It is feasible that phages with larger genomes have an extended host range, as they enclose more space to encode arrays of genes that could counteract host defenses. However, one should consider that temperate phages may have a broader host range than observed by productive infection assays. The detection of hosts, in which these phages pursue a lysogenic infection cycle, will expand the here unveiled host-range breath. In conclusion, by sequencing 40 CoNS staphylococcal phages from the same environmental niche, we greatly extend the spectrum of genome diversity. We demonstrate that phages from diverse taxonomic groups infect bacteria from numerous species, ecosystems, and drug resistant phenotypes within the genus *Staphylococcus.*

Phages from diverse taxonomic groups encapsidate foreign genetic material

In this study, we showed the existence of an expansive network among bacteria of different species mediated by phages. Ultimately, we appraised those phages' potential to incorporate

foreign genetic material. For this, we transformed a natural *S. sciuri* plasmid pUR2865[34] (3.83 kb) conferring chloramphenicol resistance into *S. epidermidis* S414. This strain was chosen as donor, as it was infected by most sequenced phages (26) and by members of the *Sipho-* and *Herelleviridae*. We propagated those phages on *S. epidermidis* S414/pUR2865 and quantified the encapsidated plasmid pUR2865 by qPCR. In addition, generalize transducing staphylococcal phage 80α and myovirus phage K were propagated on *S. aureus* RN4220/pUR2865. The removal of contaminating non-encapsulated DNA was verified using controls as established in[35]. Plasmid numbers ranged from $1.3x10^1$ to $1.6x10^6$ copies/ng phage DNA with high variations between phage samples. Using the detected copy numbers, we estimated the frequency of transducing particle formation. We assumed, that transducing particles consist of plasmid multimers only[36], and that as many base pairs of plasmid DNA are incorporated as the respective phage genome length. Figure 7a summarizes the differences in frequencies of phage transducing particles monitored per phage sample. The frequencies of transducing particles harbouring the plasmid indicate that one out of $1.5x10^2$ to maximal $2x10^7$ phages package foreign genetic material. We expected high plasmid incorporation rates for phage 80α and for the small siphoviruses, as transduction ability for those phages is generally accepted[37]. To our knowledge, there is only one report of a generalize transducing staphylococcal myovirus[26]. Strikingly, with our model, phage 80α showed comparable frequencies ($5x10^{-6}$ to $7x10^{-8}$) of transducing particles to the here characterized myoviruses. In contrast, the small siphoviruses isolated from bacterial lysogens, and one large siphovirus (PG-2021_46), showed particularly high frequencies between $6.6x10^{-3}$ and $1.6x10^{-5}$. These suggest a more targeted packaging approach of foreign genetic material. Thus, we assessed the phage genome termini, which reflect its DNA packaging mechanism (Table 2, Figure 7b). Interestingly, in several cases, predicted packaging mechanisms did not correlate with phage morphology, and we find high encapsidation frequencies for phages with other packaging mechanisms than the previously found transducing *pac*[38,39] or *cos*[40,41] phages. Yet, a *pac* mechanism is likely for the four induced small siphoviruses with high encapsidation rates, as PhageTerm predicted terminally redundant and circularly permuted genome ends. However, due to a low statistical signal, a definitive confirmation was not obtained.
Our results confirm that plasmid-borne genetic material can be used by phages for mobilization. Furthermore, we demonstrate that multiple phages from diverse taxonomic groups package foreign genetic material, albeit at various frequencies.
These data impose great potential for phage-mediated genetic transfer among bacteria, supported by the fact that phages are involved in far more numerous microbial connections than previously assumed.

Conclusion

Earlier studies have addressed the staphylococcal phage host range to predict the therapeutic fitness of phages or their impact on staphylococcal host diversity. However, phage-host arrays were limited to a small number of species[42]. In fact, most studies focused on phages from *S. aureus*[25,43-46], and only a few included phages infecting CoNS[47-50]. This restricted variety impeded broad conclusions and lead to an underestimated breath of host range for staphylococcal phages. Our data contains an unprecedented diversity, as it comprises almost 12,000 separate attempts to infect 123 hosts from 32 species with 94 different staphylococcal phage isolates. Using this phage-bacteria interaction matrix, we provide evidence that a broad host range is a dominant trait among staphylococcal phages. The ability to infect strains across the species barrier and hosts from different ecological and clinical backgrounds was not restricted to a specific phage group. On the contrary, phages with both myo- and siphoviral morphology as well as temperate and virulent lifestyle presented this trait. Our findings challenge the notion of a strong species tropism within the genus *Staphylococcus*[20,32] and confront the assumption that differences in WTA structure restrict phage infection across species. We suggest that WTAs structures are not that evolutionary divergent; phages bind to alternative, more conserved receptors on the bacterial cell wall; or phages encode multiple receptor binding proteins.

While the infection of a broad spectrum of hosts is desirable for phages in therapy, it simultaneously implies opportunities to transfer genetic material. Indeed, phage mediated horizontal gene transfer is considered to be one of the primary driving forces for the spread of antimicrobial resistance in staphylococci[16]. However, it is thought to occur rarely, and primarily within species due to estimated narrow host ranges[51]. Using a bipartite network, we demonstrate that multiple phages are shared between antimicrobial resistant and susceptible hosts, and that each drug resistant host in this network is, on average, connected to 16 drug susceptible neighbors. The many connections and routes confirm the potential role of phages in the mobilization and dispersal of genetic material. Nevertheless, transduction ability has, so far, only been awarded to some staphylococcal phages[37]. On these grounds, we quantified bacterial DNA encapsidation rates for 19 myoviruses and 9 siphoviruses from this network. We detected packaged plasmid DNA in all assessed phages, confirming this competence as widespread among staphylococcal phages[26,37,39]. Our data indicate that one phage particle out of every hundred to maximal 10^7 phage particle is transducing. However, those numbers do not necessarily reflect the frequency of generalized transduction due to the following reasoning. We propose that within phage transduction one must acknowledge two main bottlenecks. First, the capability of phages to incorporate foreign DNA and at which frequency transducing particles are being formed. This is dependent on individual phage characteristics, and on type and location of the bacterial cargo DNA within the host. Second, the delivery and

expression of the cargo DNA in the recipient bacteria. This can highly differ between strains, as it is mostly depending on the bacterial “immune system” such as restriction modification systems and CRISPR-Cas[10]. In simplified models, studies propose that transduction efficiencies, thus the successfully delivery and expression of cargo DNA in a recipient bacterium, is approximately 3 % [38,39]. To this regard, upcoming studies will determine the ability of the here detected transducing particles to spread the drug resistance element across this unique network.

In conclusion, this study reveals an expansive interspecies communication network and place phages as central mediators for bacterial connectivity. Our findings support the speculated interspecies horizontal transfer of adaptive genetic material by phages and exemplify the impact of phage populations on the evolution of human pathogens.

Methods

Wastewater sampling

Water samples were collected at the wastewater treatment plant in Au (Zurich, Switzerland) on July 24th, 2018. The WWTP Rietliau receives 7 - 30 million L of wastewater a day and processes it within 24 hours. The treated water, of which half is filtered through a 0.035 µm membrane, is directly released into Lake Zurich. Samples (2.5 L each) were taken at the entrance after the mechanical clearance and from the effluent. Both inlet and outlet samples were centrifuged at 10,000 rpm for 30 minutes at 4 °C. The supernatants were 0.22 µm PES sterile filtered and kept at 4 °C for phage enrichment. The bacterial pellets were suspended in 20 ml 0.85 % NaCl and held at -20 °C for prophage induction.

Bacterial Strains, Culture Conditions and Plaque Assay

Bacterial strains for this study[44,52-85] were seeded on tryptic soy agar (TSA, 2 % agar and 30 g/L tryptic soy broth (TSB)) and grown in TSB overnight at 37 °C. Plaque assays were carried out using LC agar as top agar (10 g/L casein peptone, 5 g/L yeast extract, 128 mM NaCl, 55.5 mM glucose, 2mM $MgSO_4$, 10 mM $CaCl_2$, 0.4 % agar), and TSA as bottom agar. Phages (10 µL, serially diluted) were mixed with bacterial hosts in molten soft agar (47 °C), plated, and incubated overnight before quantification. For spot assays, bacteria were inoculated into molten soft agar (47 °C), plated, and phage concentrates (5 µL, serially diluted) were dropped onto.

Enrichment Cocktail Constitution

Five cocktails were generated to enrich staphylococcal phages from wastewater. Staphylococcal strains for each cocktail were selected to produce a diverse community and combined either randomly (cocktail A), or according to their origin (cocktail B: animal related strains; cocktail C and D: environmental isolated strains; cocktail E: lab strains). To assure growth harmony for each bacterium within a cocktail, strains with cross-infective prophages or bacteriocin producers were excluded. For this, all selected strains were induced using Mitomycin C and UV irradiation (protocol adapted from[86]). Briefly, 50 µL of a fresh overnight culture was inoculated in 5 mL TSB and incubated on a shaker for 2 hours at 37 °C. The initial absorbance was measured at OD_{600}. Mitomycin C was added to a final concentration of 0.5 µg/mL, and bacterial suspensions were shaken at 37 °C. For UV irradiation, cells were centrifuged at 6,000 x g for 10 minutes at room temperature. The pellet was resuspended in 5 mL 0.1 M $MgSO_4$ and irradiated with UV-Light (2400 µJ/cm^2). After irradiation, cells were transferred to double strength TSB, protected from light, and incubated on a shaker at 37 °C. The absorbance of both UV and Mitomycin C induced strains was then measured every hour for 6 hours or until a decrease of the OD_{600} was observed. The bacterial cultures were then

centrifuged at 3,000 x g for 12 minutes at 4 °C, the supernatant 0.22 µm sterile filtered, and stored at 4 °C. For all induction experiments, *S. aureus* Newman served as a positive control, as it contains three inducible prophages that lyse *S. aureus* RN4220[87]. Spot assays were performed to assess the presence of cross-reactive phages that interfere with the growth of strains within a cocktail.

The radial streak method was applied to determine whether cocktail members restrain the growth of others by the production of bacteriocins or other extracellular antimicrobial compounds (protocol adapted from[88]). In short, the area of a small circle was inoculated with a 0.5 McFarland bacterial suspension of each cocktail candidate member in the center of a fresh plate. The plates were incubated at 37 °C for 24 hours, and all remaining members of the respective cocktail (0.5 McFarland) were then radially streaked from the border of the dish to the circle area. If the central bacterial strain provoked a zone of growth inhibition after a second incubation, it was excluded.

Phage Enrichment and Isolation

Inlet and outlet phage suspensions were enriched for staphylococcal phages using the five constituted enrichment cocktails independently. For each cocktail and sample, 80 mL of the viral suspension was supplemented with 20 mL 5 x TSB and 100 µL of a fresh overnight culture of every cocktail member. The ten suspensions were then incubated overnight at 37 °C. After this first round of enrichment, viral suspensions were centrifuged at 10,000 rpm for 30 minutes at 4 °C, and the supernatants 0.45 µm PES sterile filtered. For the second enrichment, 20 ml of 5 x TSB and 100 µl of a fresh overnight culture of the same cocktail members were added anew and processed as described above. The enrichment process was repeated for a total of three rounds. Staphylococcal phages were detected by spotting 10 µl of the enriched viral suspensions on a bacterial lawn of each enrichment host, and plates were incubated overnight at 37 °C. If a zone of lysis or individual plaques were visible the next day, a plaque assay was performed with serially diluted phage suspensions. Plates with single lysis plaques were examined for different plaque morphologies, and a maximum of three were picked for phage purification for each plate. Phages were purified by repeatedly plating and picking individual plaques for three rounds.

For prophage induction and isolation, bacterial pellets frozen from wastewater were thawed and resuspended in 20 ml double strength TSB supplemented with 6.5 % NaCl for staphylococcal enrichment. After overnight incubation, 10 ml of each enrichment was added to 490 ml TSB, and the initial absorbance (OD_{600}) was measured. Cells were grown until an OD_{600} of 0.5, and the sample split for the induction with Mitomycin C or UV irradiation. For Mitomycin C induction, a final concentration of 1 µg/mL was added, and the suspension was incubated at 37 °C for 6 hours. For UV irradiation, cells were centrifuged at 6,000 x g for 10

minutes and the pellet resuspended in 125 mL 0.1 M $MgSO_4$. This resuspension was irradiated (4400 µJ/cm^2), transferred to 125 ml double strength TSB, protected from light, and was incubated for 6 hours at 37 °C. Finally, induced samples were centrifuged at 10,000 x g for 15 minutes at 4 °C, the supernatants 0.22 µm PES sterile filtrated, and stored at 4 °C. For temperate phage detection, serially diluted phage suspensions were dropped on all hosts selected for host range determination (Hosts in Supplementary Table 10). If either a zone of lysis or individual plaques were visible after overnight incubation, phages were picked and purified as described above.

Phage Host Range Determination

Phage host ranges were assessed on 123 strains (32 species) that originated from human (40), veterinary (53), or environmental settings (23) harboring a multidrug resistant (35), resistant (49), or antibiotic susceptible phenotype (40). The hosts were chosen to represent a diverse community of both CoNS (68) and CoPS (49), as well as other Gram-positive bacteria (6) (Supplementary Table 10). For the classification of multidrug resistant strains, bacteria resistant to three or more antibiotic families were considered multi-drug resistant, whereas the Macrolide-lincosamide-streptogramin B (MLS_B) resistance phenotype was classified as one family. Staphylococcal strains with an unknown coagulase phenotype were assessed for coagulase production using Staph Rapid Latex Test Kit (Brunelli, #271060). Each phage lysate was spotted (5 µl) in duplicates at five concentrations (10^8-10^4 pfu/ml) onto those selected hosts. If single lysis plaques appeared in any dilution after overnight incubation, the strain was considered susceptible to the respective phage, and an infection event was reported. Lysis from without (LFW) events, where a bacterial lysis halo without single visible plaques appears, were additionally reported but not considered as infection. *Staphylococcus* phage K propagated on *S. aureus* PSK ATCC 19685 was used as a reference for all host range assays. Phages with equal host ranges on all 123 hosts were clustered, and further characterizations were continued with one selected phage per cluster.

Phage Propagation

Phages were produced using the double-agar-layer method and washed off 20 to 80 semi-confluent lysis plates using SM buffer (200 mM sodium chloride, 10 mM $MgSO_4$, 50 mM tris, and 0.01 % gelatin, pH 7.4) and agitation for 4 hours (20 rpm). The phage lysates were collected, and cellular debris or agar remnants were removed by centrifugation at 5,000 x g for 10 minutes at 4 °C. The supernatant was 0.22 µm sterile filtrated. Phage particles were precipitated with 7 % PEG_{8000} supplemented with 1 M NaCl in ice water for two days. The precipitated phages were collected by centrifugation at 10,000 x g for 20 minutes at 4 °C, and pellets were dissolved in 8 mL SM buffer. Phages were purified by CsCl ultracentrifugation. Briefly, the density of each phage suspension was adjusted 1.15 g/mL using CsCl and added

on top of a three-layer (1.7, 1.5, and 1.35) CsCl density gradient. The gradient was centrifuged at 82,000 x g for 2 hours at 10 °C, and the phages were collected between the 1.35 and 1.5 density layers. All purified phages were dialyzed overnight at 4 °C in 4 L SM buffer (50 kDa cut off) under gentle magnetic stirring.

Phage DNA Extraction

Phage DNA was extracted using the phenol/chloroform DNA extraction method. In short, 640 µL of propagated phage lysate (> 10^{10} pfu/mL) were treated with 10 U DNase I for 1 hour at 37 °C, and the enzyme heat-inactivated for 10 minutes at 65 °C in the presence of 20 mM EDTA. Proteinase K was added to a final concentration of 100 µg/ml, the sample vortexed and incubated for 1 hour at 50 °C, 300 rpm. Next, one volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added, the sample centrifuged for 13'000 x g for 15 minutes, and the aqueous layer extracted. This step was repeated with 1 volume chloroform:isoamyl alcohol (24:1). DNA was precipitated by adding 50 µl 5 M NaCl and 0.7 volumes of isopropanol. The next day, the DNA was pelleted with 13'000 x g for 20 minutes at 4 °C, and the pellet washed twice with ice-cold 70 % EtOH. DNA was resuspended in 50 µl 10 mM Tris (pH = 8.0), and the concentration was measured using Qubit.

Electron Microscopy

Propagated phages (≥ 10^9 pfu/ml, 8 µl) were let absorb to negatively discharged (45 seconds, $3x10^{-1}$ mbar, 25 mA) carbon-coated copper grids (Quantifoil) for one minute. Grids were washed twice in pure water and adsorbed particles negatively stained for 20 seconds with 2 % uranyl acetate or phosphotungstic acid. They were observed at 100 kV on a Hitachi HT 7700 scope equipped with an AMT XR81B Peltier cooled CCD camera (8M pixel).

Genome Library Preparation, Sequencing and Bioinformatics

Forty phages were selected for whole genome sequencing. Precedence was given to phages obtained from the WWTP outlet and bacterial lysogens, and later those that infected hosts from diverse ecosystems and with different drug resistant phenotypes. Phage genomes were Illumina sequenced if genomic DNA yields were <1 µg. For Illumina sequencing, multiplexed libraries were prepared using the Illumina Trueseq Nano library prep according to manufactures' instructions. Phage DNA was paired-end sequenced with 0.5 million reads (150 bp/read) using the MiSeq sequencer. Raw reads were trimmed with Trimmomatic[89] in default settings and assembled using SPAdes[90] in careful mode. For Pacbio sequencing, gDNA (ca. 1 µg) was mechanically sheared to the average size distribution of 8-10 kb, using a Covaris gTube (Covaris p/n 520079). Multiplex libraries were prepared using the SMRTBell™ Barcoded Adapter Complete Prep Kit - 96, following the manufacturer's instructions (100-514-900, Pacific Biosciences). Tagged libraries were sequenced in a 1M SMART Cell with PacBio

Sequel. De-multiplexed reads were assembled using the Hierarchical Genome Assembly Process[91] (HGAP4, SMRT Link v8.0.0). When needed, Sanger sequencing was used to close gaps in the assembled genomes. Open reading frames (ORFs) were predicted with PHANOTATE[92] and annotated using multiPhATE[93] with blastn against the NCBI virus, blastp against pVOGs[94], PhAnToMe, and NCBI virus, and jackhammer against the pVOGS database. Potential tRNAs in phage genomes were predicted using tRNAScan-SE v2.0.5[95]. Phage termini were predicted using PhageTerm[96].

Phylogenetic Analysis

Biopython 32 package was used within the conda environment to retrieve fully sequenced staphylococcal phage genomes deposited at GenBank as of June 2020 (n = 292)[19]. Unverified cRNA or partial phage genomes were excluded from the analysis. The closest relative on NCBI was determined by average nucleotide identity (ANI) values as in[97,98]. Distances between genomic sequences for phylogenomic analysis were calculated as described in[99,100], and the tree visual represented in iTOL[101].

Network Analysis

The network analysis was based on the host range matrix consisting of 123 bacterial hosts and 94 phages isolated from wastewater. A binary incidence matrix was generated from the data in which infections are indicated as one, and no interaction is marked as zero. Phage resistant hosts (n = 63/123) were removed, and a bipartite network was generated using the R package igraph[102]. In this network, phage permissive bacteria (60) and respective phages (94) were represented as nodes where an edge between a bacterial and phage node indicates phage infection. This network was further collapsed into a bipartite projection, in which only bacteria are represented as nodes and phages as edges connecting two bacterial nodes. The number of shared phages between two nodes was assigned as an edge attribute. Best connected hosts were identified by the highest number of shared phages. The mean number of shared phages was calculated by averaging the values for all possible host pairs (M=60x59/2=1770) in the bipartite projection. Host pairs with no shared phages (740) were included in the average with a value of zero. Direct neighbors (degree of a node) were counted as the sum of all nodes that are connected by an edge to a specified node. For a subset neighbor count, only neighbors with a specific attribute like resistance profile, origin, or species affiliation were considered.

Analysis of Nestedness and Modularity

Modularity and nestedness were calculated based on the generated biadjacency matrix. All bacteria resistant to phage infections (n = 63/123) were removed from the dataset, resulting in a 60 bacteria x 94 phage matrix. Modularity (Q) was calculated with the Ipbrim R package[103]

using the findModules function with 100 iterations. Nestedness was measured with the nestednodf function of the vegan package[104] in R. The null mode method "r00" was chosen as it preserves the matrix size and number of interactions. Statistical significance was evaluated using the oecosimu function with 100 simulations (one-sided testing with statistic assumed greater than simulated values).

Encapsidation Rates

A subset of sequenced staphylococcal phages were assessed for their ability to encapsidate foreign genetic material. For this, the donor *S. epidermidis* S414, susceptible to most sequenced staphylococcal phages (26 out of 40) was selected. A small, natural occurring *S. sciuri* plasmid pUR2865 (3.83 kb) conferring resistance to chloramphenicol (*cat*$_{pC221}$) was chosen as genetic marker for encapsidation. The plasmid was transformed into *S. epidermidis* S414 and *S. aureus* RN4220 using methods established in [105]. Sequenced phages infecting *S. epidermidis* S414/pUR2865 were then propagated on this strain. Equally, staphylococcal phage K and phage 80α were propagated on *S. aureus* RN4220/pUR2865. Phage particles were washed off three semi-confluent lysis plates using SM buffer (see Phage Propagation) and supernatants were 0.22 µm sterile filtrated. Phage lysates were purified using CsCl density gradient centrifugation and dialysed (see Phage Propagation). Samples (620 µl) were treated with 100 Units DNase I and phage encapsidated DNA was extracted (see Phage DNA Extraction). DNA concentrations were measured in duplicates using Qubit Fluorometric Quantification (Thermo Fisher Scientific). Copy numbers of the chloramphenicol resistance marker *cat*$_{pC221}$ (pUR2865) were quantified by Taqman qPCR in triplicates using the Roche LightCycler480 system. Primers were as follows: catpC221-fw (GTAACAATAGCAGCTTTTTATTGCCT), catpC221-rv (TAAATAATGAAGCATGGTAACCATCAC) and catpC221-probe (AGCATGATGAAGCTGTAAGGCAACTGGTAT) (product length, 132 bp). Each reaction mixture (20 µl) contained 10 µl SensiFAST Probe No-ROX Kit 2X (Labgene Scientific), 0.25 µM Probe, 0.9 µM of each primer and 1 µl (3 ng) of the extracted phage DNA. The standards for *cat*$_{pC221}$ ranged from 10^7 copies/µl in 10-fold dilution to 10^1 copies/µl, respectively. Initial polymerase activation at 95 °C for 5 minutes was followed by 45 cycles of denaturation at 95 °C for 10 seconds, and amplification at 58 °C for 20 seconds. To exclude the possibility of non-encapsidated DNA contaminants, several controls were added. For this, the absence of non-packaged DNA, the DNase I activity and inactivation were tested as introduced in[35]. Furthermore, the absence of any contaminating extracellular DNA was verified through a no-phage control. These samples were treated equally to regular phage samples, and the absence of extracellular plasmid DNA was verified by qPCR after DNaseI treatment. Encapsidation frequencies were calculated as follows: First, detected copy numbers for pUR2865 were normalized to 1 ng of DNA (A). Next, the number of respective phage genome

copies in 1 ng DNA (B) was calculated using the following formula[38]: copies = ((1ng x NA x10^9)/Mr), where Mr = size of the phage genomic DNA (bp) multiplied by normalized weight of nucleotide base (650 Da), and NA is the Avogadro constant. Lastly, encapsidation frequencies were calculated using the formula: EF=(A/(BxC)), where C indicates the number of plasmids that can be packaged into each respective phage capsid[36] (phage genome (bp) / plasmid genome (bp)).

Statistics

All test statistics were calculated with R, using the base package stats. For data manipulation and plotting, dplyr[106] and ggplot2[107] were used. All scripts are available as an R Markdown upon request. If not otherwise indicated, an average value is always displayed with its corresponding standard derivation.

Data Availability

All bacteriophages are available upon request. Sequenced phage genomes are available under the bioproject PRJEB42698 in the European Nucleotide Archive, Sample 1-40.

Abbreviations

°C	degrees Celsius
µJ	microjoule
µL	microliter
µm	micrometre
ANI	average nucleotide identity
cm^2	square centimetre
CoNS	coagulase negative *Staphylococci*
CoPS	coagulase positive *Staphylococci*
CsCl	cesium chloride
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
g	gram
HCL	hydrochloric acid
HGAP4	hierarchical genome assembly process
L	liter
M	molar
$MgSO_4$	magnesium sulfate
MIC	mean inhibitory concentration
ml	millilitre
MLS_B	Macrolide-linocosamide-streptogramin B (MLSB) resistance phenotype
mM	milimolar
NaCl	sodium chloride
ng	nanogram
ORFs	open reading frames
PEG	polyethylene glycol
PES	polyethersulfone
rpm	revolutions per minute
TSA	tryptic soy agar
TSB	tryptic soy broth
WTA	wall teichoic acid

Acknowledgements

This work was supported by the Swiss National Science Foundation NFP72 “Antimicrobial Resistance” Project No. 167090, by the European Union’s Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Sklodowska-Curie Grant Agreement No. 659314, and by the ETH Career Seed Grant Project SEED-01 18-1.

We kindly thank Carmen Torres, University of La Rioja, Spain; Antonella Demarta, SUPSI-Laboratory of Applied Microbiology, Switzerland; Vincent Perreten, University of Bern, Switzerland, for providing bacterial strains used in this study. We thank Daniel Fehlmann, WWTP Wädenswil, for providing wastewater samples, determining the chemiophysical parameters, and the stimulating exchange. Further, we would like to thank Hugo Oliveira for sharing python codes to extract staphylococcal phages from NCBI. We thank Jochen Klumpp and Stefan Handschin for their electron microscopy expert advice. Lastly, we thank Andrea Hauser, Jose Manuel Haro-Moreno, Diana Gutierrez, Jonas Fernbach and Christian Röhrig for helpful and expert discussions.

Authors Contribution

P.C.G. guided and analyzed all experiments and wrote the manuscript. E.G.S conceived the study, guided experiments, and contributed to writing of the manuscript. T.E. and P.C.G performed the bioinformatic, network and encapsidation analysis. T.E. wrote all R scripts. D.L., V.B., N.R., and N.E. contributed to phage isolation. A.N. and P.C.G. propagated sequenced phages. N.R., N.E. and P.C.G performed the host range assays. E.K. and V.B. established bacterial enrichment cocktails. F.H.C. analyzed sequenced phage genomes phylogenetically and supported bioinformatic analysis. M.J.L provided conceptual input, partial funding, and corrected the manuscript. All authors read and approved the final manuscript.

Competing Interests

The authors declare no competing interests.

Tables

Table 1

General properties of the phage-bacteria interaction network.

Network		
	# host species (strains)	32 (123)
	# phages	94
	# interactions (I)	1135
	Size (M)	11562
	Connectance (C = I/M)	0.098
Hosts		
	Infected species (strains)	27 (60)
	Mean (± sd) phage infection per strain	9.2 ± 15
	Mean (± sd) phage infection per species	22.3 ± 22.5
	Maximal phage infections per strain	65
	% infections on CoNS	89%
	% infections on CoPS	11%
Phages		
	Maximal species (strains) infection per phage	32 (18)
	Mean (± sd) species infections per phage	7.8 ± 3.7
	Mean (± sd) strain infections per phage	12.10 ± 5.4

692 **Table 2**

693 Characteristics of the 40 sequenced staphylococcal phage genomes.

Phage[1]	ORI[2]	MO[3]	HR[4]	Propagation	Genome (bp)	Termini[5]	%GC	ORFs	tRNA	Closest Relative	%ANI
1	I	M	12/7	*S. epidermidis*	143'764	DTR (long)	27.96	241	-	Twillingate	98.01
2	I	M	13/9	*S. vitulinus*	142'223	DTR (long)	30.85	244	1	vB_Sau_Clo6	92.14
4	O	S	3/3	*S. sciuri*	91'860	pac	30.61	171	-	vB_StaM_SA2	67.09
5	O	S	10/5	*S. epidermidis*	92'130	cos	29.41	174	-	6ec	95.45
8	I	M	20/14	*S. equorum*	139'709	unknown	30.94	230	-	vB_Sau_Clo6	91.42
9	I	M	25/16	*S. xylosus*	141'528	DTR (long)	30.77	244	-	vB_Sau_Clo7	92.24
10	I	M	22/15	*S. xylosus*	141'528	DTR (long)	30.77	270	-	vB_Sau_Clo8	92.12
12	I	M	15/9	*S. vitulinus*	145'091	DTR (long)	31.33	285	3	vB_SscM-2	67.88
14	I	M	14/10	*S. vitulinus*	145'090	DTR (long)	31.33	245	3	vB_SscM-4	69.95
15	I	M	18/11	*S. xylosus*	145'090	DTR (long)	31.33	285	-	vB_SscM-5	67.83
16	I	M	19/11	*S. xylosus*	141'321	DTR (long)	30.80	213	-	vB_Sau_Clo6	91.40
17	I	M	32/18	*S. xylosus*	144'971	DTR (long)	30.85	261	2	vB_Sau_Clo6	91.85
18	O	M	15/10	*S. xylosus*	138'844	DTR (long)	30.80	225	2	vB_Sau_Clo6	92.15
19	I	M	16/9	*S. vitulinus*	141'132	DTR (long)	31.25	247	2	vB_Sau_S24	88.86
22	I	M	13/8	*S. sciuri*	144'280	DTR (long)	31.33	250	-	vB_SscM-1	69.16
23	I	M	14/9	*S. vitulinus*	139'827	DTR (long)	28.00	236	-	Twillingate	97.63
27	I	M	4/4	*S. aureus*	128'279	DTR (long)	29.67	220	-	Quidividi	69.15
29	I	M	15/11	*S. epidermidis*	131'570	pac	30.89	215	1	VB_SavM_JYL01	91.34
31	I	M	10/7	*S. vitulinus*	139'439	unknown	31.59	244	-	vB_SscM-1	90.11
33	I	M	12/8	*S. vitulinus*	135'943	pac	31.67	234	-	vB_SscM-1	90.25
35	I	M	14/9	*S. xylosus*	138'653	DTR (long)	30.80	254	-	vB_Sau_Clo6	92.19
38	I	M	19/11	*S. epidermidis*	140'647	DTR (long)	30.80	260	1	vB_Sau_S24	92.47
40	I	M	13/8	*S. vitulinus*	142'875	DTR (long)	31.35	270	3	vB_SscM-2	69.35
41	I	M	13/7	*S. sciuri*	145'090	DTR (long)	31.34	244	3	vB_SscM-1	68.67
43	I	M	15/9	*S. sciuri*	145'090	DTR (long)	31.34	287	3	vB_SscM-1	67.88
46	O	S	8/5	*S. epidermidis*	86'018	DTR (short)	29.66	152	-	6ec	94.90
47	O	M	19/11	*S. xylosus*	142'885	DTR (long)	30.70	238	-	vB_Sau_Clo6	92.32
64	I	M	19/14	*S. succinus*	142'287	DTR (long)	30.89	231	1	vB_Sau_Clo6	91.52
67	I	S	3/2	*S. sciuri*	92'064	cos	30.57	189	2	vB_StaM_SA2	66.72
68	I	S	4/3	*S. sciuri*	91'947	cos	30.61	197	1	vB_StaM_SA2	66.92
74	O	S	9/5	*S. epidermidis*	85'762	pac	29.66	150	1	6ec	95.62
76	I	M	10/6	*S. vitulinus*	139'439	unknown	31.59	245	-	vB_SscM-1	90.13
84	I	M	19/12	*S. xylosus*	139'439	DTR (long)	31.59	246	2	vB_Sau_S24	92.22
86	I	M	22/14	*S. xylosus*	141'291	DTR (long)	30.76	228	2	vB_Sau_Clo6	90.43
87	I	M	21/13	*S. xylosus*	141'212	DTR (long)	30.75	239	-	vB_Sau_Clo6	92.65
88	O	S	9/4	*S. epidermidis*	92'222	DTR (long)	30.83	174	1	6ec	95.14
89	ID	S	5/1	*S. epidermidis*	43'039	unknown	35.11	77	-	IME1348_01	95.03
90	ID	S	6/2	*S. epidermidis*	44'493	unknown	34.72	65	-	IME1348_01	94.72
91	ID	S	6/2	*S. epidermidis*	42'188	unknown	34.97	61	-	IME1348_01	95.82
93	ID	S	4/1	*S. epidermidis*	43'459	unknown	34.37	79	-	SepiS-phiIPLA7	96.16

694 [1] Phages are abbreviated with their final, unique numerical identifier (PG-2021_*).
695 [2]Isolation origin (I: inlet, O: Outlet, ID: Induced). [3]Morphology (M: Myovirus, S: Siphovirus).
696 [4]Host range as number of strains/species infected. [5]Termini (DTR: direct terminal repeats).

697

Table 3

General properties of the bipartite network projection. In the bipartite network projection, hosts are nodes and the number of shared phages are weighted edges between nodes.

Bipartite Network Projection	
# of hosts (strains, H)	60
# of host species	27
# phages (P)	93
Size/Possible Interactions (M = H x (H-1) / 2)	1770
Number of interactions (I)	1030
Connectance (C = I /M)	0.58
Shared phages	
Mean (± sd) between two hosts	4.3 ± 7.9
Maximum between two host	58
Mean (± sd) between species	4.1 ± 7.5
Mean (± sd) within species	8.3 ± 11.4
Mean (± sd) between environments	4.1 ± 7.6
Mean (± sd) within environments	4.7 ± 8.4
Mean (± sd) between drug susceptible and resistant	4.3 ± 8.0
Neighbors	
Mean (± sd) neighbors per strain	34.3 ± 13.6
Maximum neighbors per strain	56
Mean (± sd) species neighbors per strain	17.4 ± 5.2
Maximum species neighbors per strain	25
Mean (± sd) neighbors per strain from other environments	23.1 ± 9.9
Mean (± sd) neighbors per strain from the same environment	11.2 ± 7.0
Mean (± sd) of drug susceptible neighbors for each drug resistant host	16.0 ± 6.7

Supplementary Table 1

Bacterial host selected for the enrichment cocktail constitution and phage isolation.

Supplementary Table 2

Summary of the constitution and phage isolation efficiency for each enrichment cocktail.

Supplementary Table 3

Compilation of all enrichment hosts and their respective efficiency in phage isolation.

Supplementary Table 4

The number of isolated phages for enrichment species, and their corresponding number of successful and unsuccessful enrichment strains.

Supplementary Table 5

Compilation of isolation hosts for all induced phages, and their respective efficiency.

Supplementary Table 6

Compilation over all isolated phages, their isolation origin and hosts, as well as the corresponding enrichment cocktail. Phages with equal cluster numbers (column F) had identical host ranges on 123 different bacteria.

Supplementary Table 7

Summary of all phage isolation and discrimination advances on each staphylococcal species.

Supplementary Table 8

Biadjacency matrix of the phage-bacterium network with characteristics and phenotypes of all challenged bacteria.

Supplementary Table 9

Corresponding enrichment cocktail for each phage cluster.

Supplementary Table 10

Isolation frequency and hosts for all isolated phages.

Supplementary Table 11

Biadjacency matrix of the phage-bacterium network with characteristics and phenotypes of all isolated phages, including phage K.

Supplementary Table 12

Taxonomic diversity of species and strains detected in each module after modularity sorting of the phage-bacteria interaction matrix.

Supplementary Table 13

Strain abundance and phage permissiveness for each bacterial species included in the host array.

Supplementary Table 14

Host range characteristics for all 94 staphylococcal phages concerning infection of strain from different ecosystems and antimicrobial resistance phenotypes.

Figures

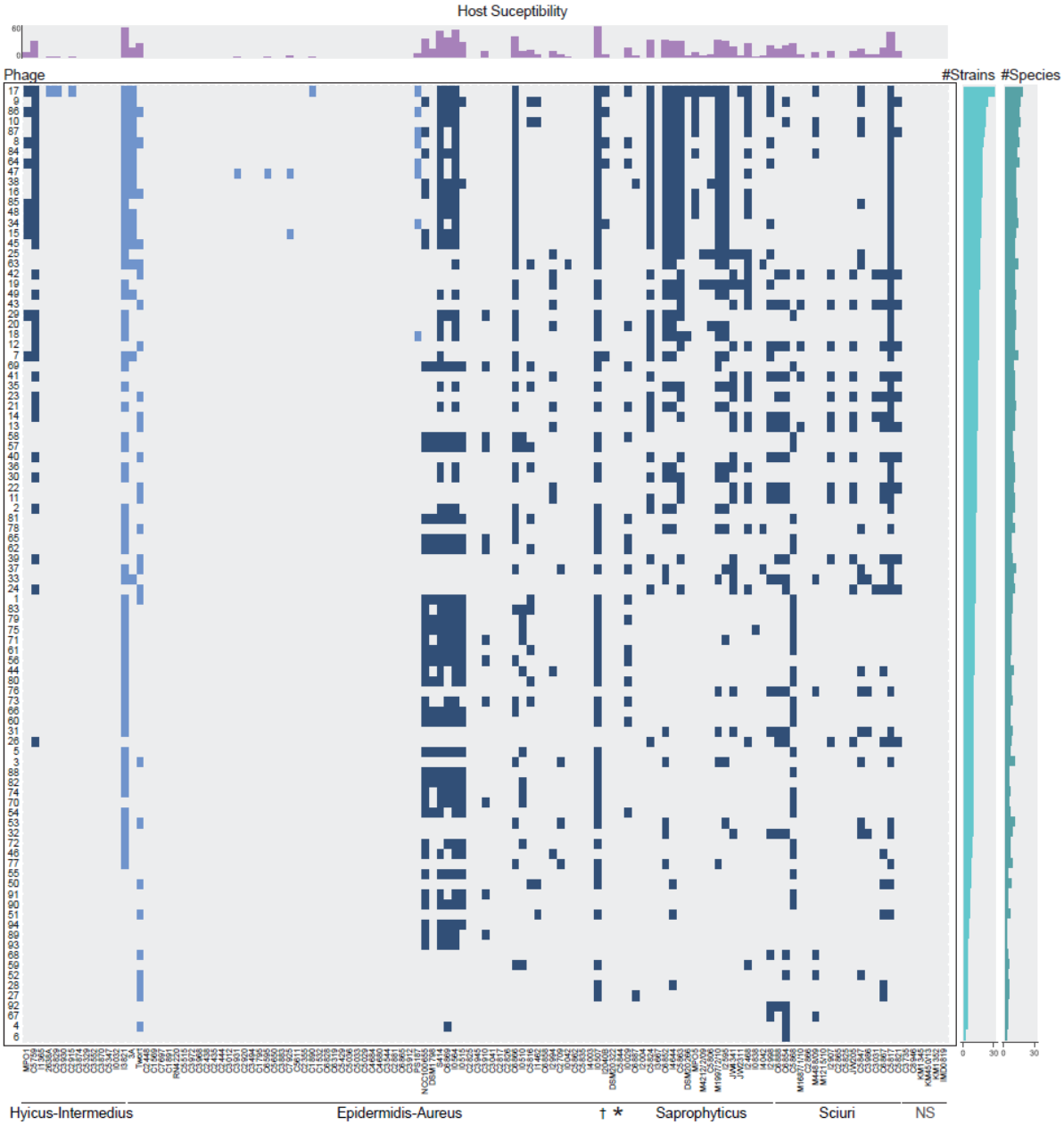


Fig 1

A staphylococcal phage-bacteria incidence matrix. Bacterial lawns of 123 hosts from 32 species, were challenged with 94 different staphylococcal phages from wastewater and phage K. Phages on the y-axis are sorted from broad host range to narrow. Bacterial hosts in columns are sorted after cluster-groups and subdivided species as established in [17,18]. †: Species-group Auricularis. *: cluster-group Simulans. NS: Non-S*taphylococcus* hosts. Each blue-colored square of the incidence matrix corresponds to a phage-host infection where single plaques were visible. Squares in dark blue indicate infections on CoNS and squares in light blue on CoPS. The phage permissiveness for each host is indicated in the host susceptibility bar chart

on top of the incidence matrix, which represents the number of phages infecting a strain. The two bar charts on the right indicate the total number of strains (# strains, left) and species (# species, right) a phage infected. The incidence matrix has a diameter of six and a density of 0.1 (=1135/11562). Phages are abbreviated with their final unique numerical identifier (PG-2021_*).

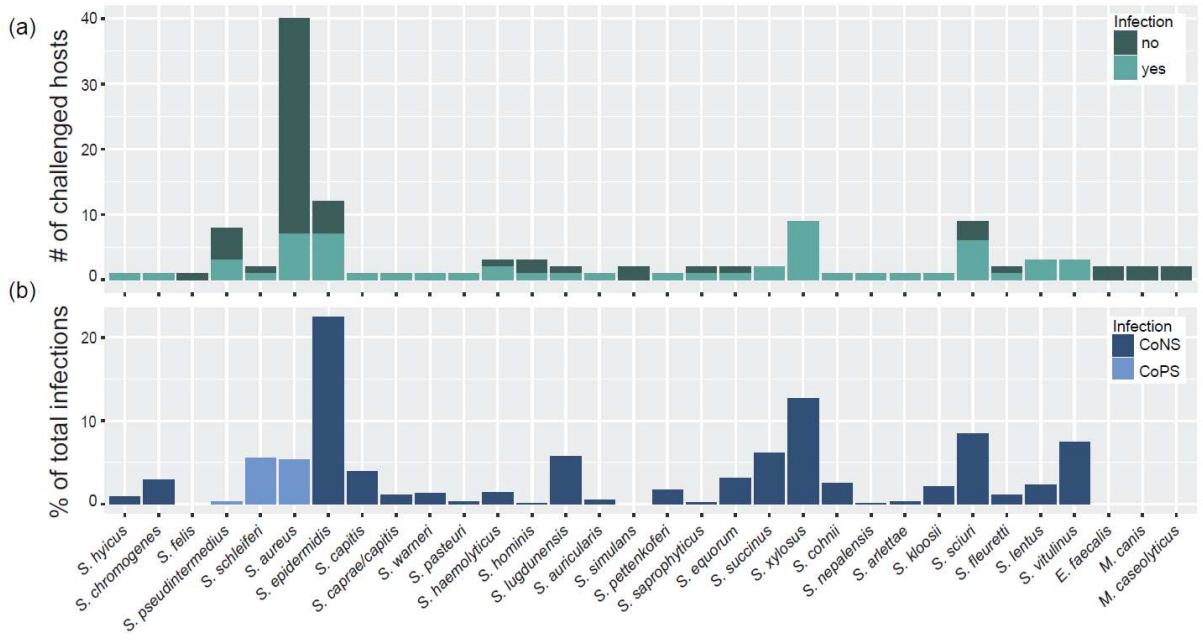


Fig 2

Phage infections on staphylococcal species. Species challenged in the phage-bacteria interaction matrix are shown on the x-axis and sorted after the established *Staphylococcus* species groups[18]. (a) For each species, the number of phage resistant and susceptible strains are depicted. (b) Phage infections on each respective species was plotted as a percentage of the total infections detected in the interaction matrix.

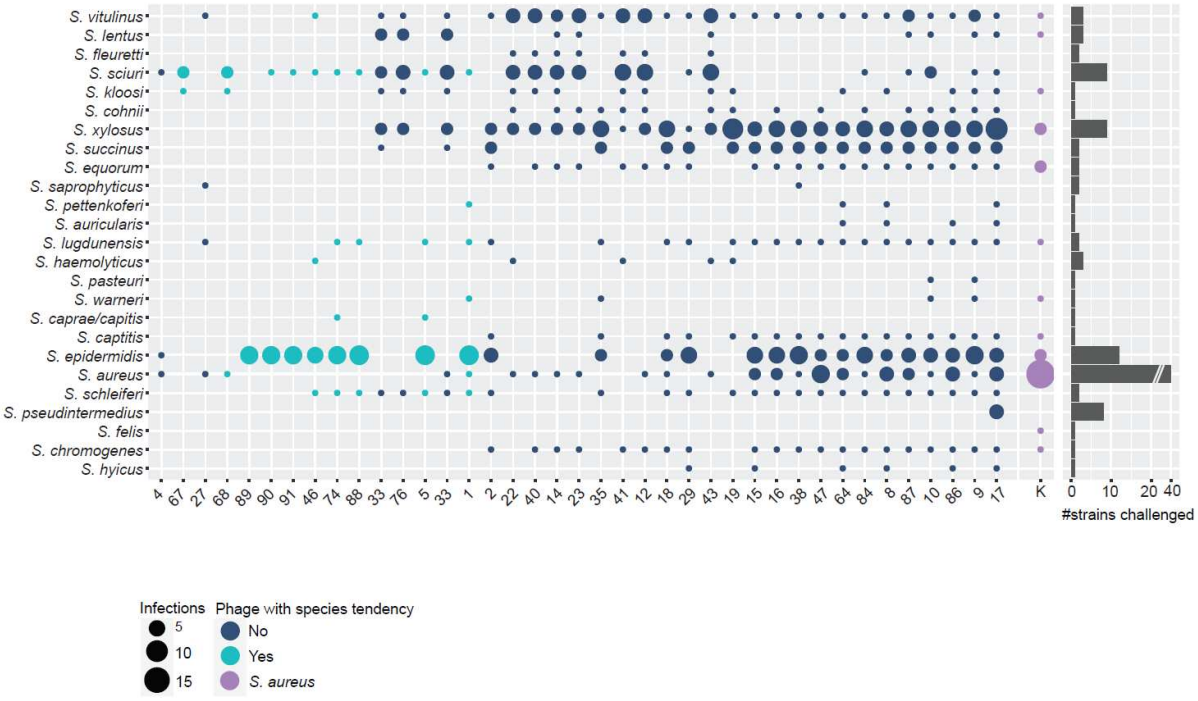


Fig 3

Illustration of the host ranges collapsed on the species level for the 40 sequenced phages. Phages on the x-axis are sorted from narrow (left) to broad host range (right). Species on the y-axis are sorted after phylogenetic relationship in species groups[18]. A phage host range is depicted as a column, where infection of a staphylococcal species is illustrated using circles. For each respective species, the area of the circle is scaled according to the number of strains a phage can replicate on (scale: 1-15). The total number of strains challenged per species is depicted in the bar-chart on the right. Host ranges on this host array are colored as follows: Phage with species tendency (≥ 50 % of all infections on a single species) in turquoise; phages with no clear species tendency in dark blue; polyvalent phage K in violet. Phages are abbreviated with their final unique numerical identifier (PG-2021_*).

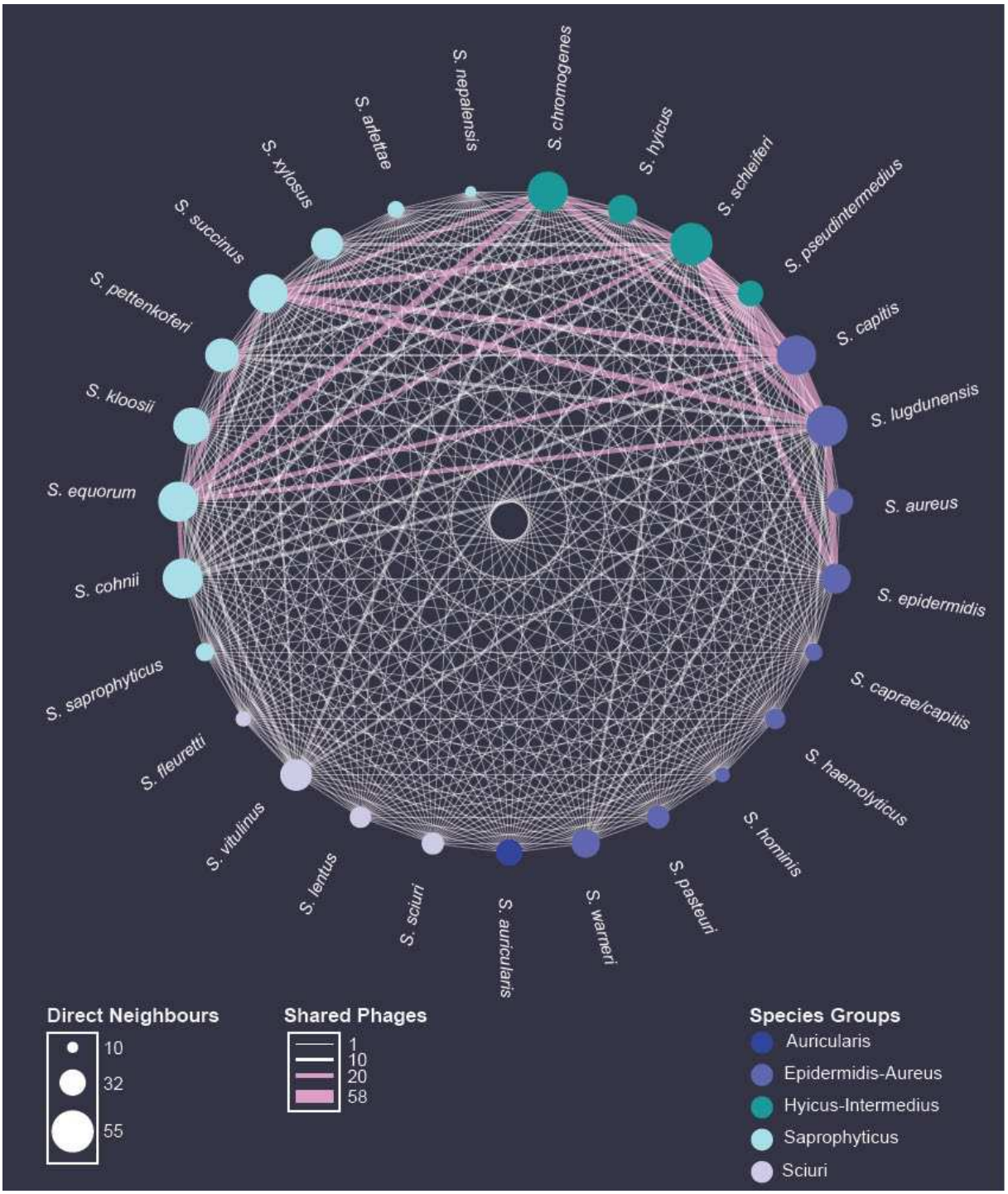


Fig 4

Species network with phages as coupling links. Staphylococcal host species are represented as nodes and sorted after cluster affiliation[18]. The area of each node directly correlates with the average number of strain neighbors a species is connected to. The number of shared phages between species is represented as weighted edges. If >20 phages are shared between two staphylococcal species, edges are colored in pink.

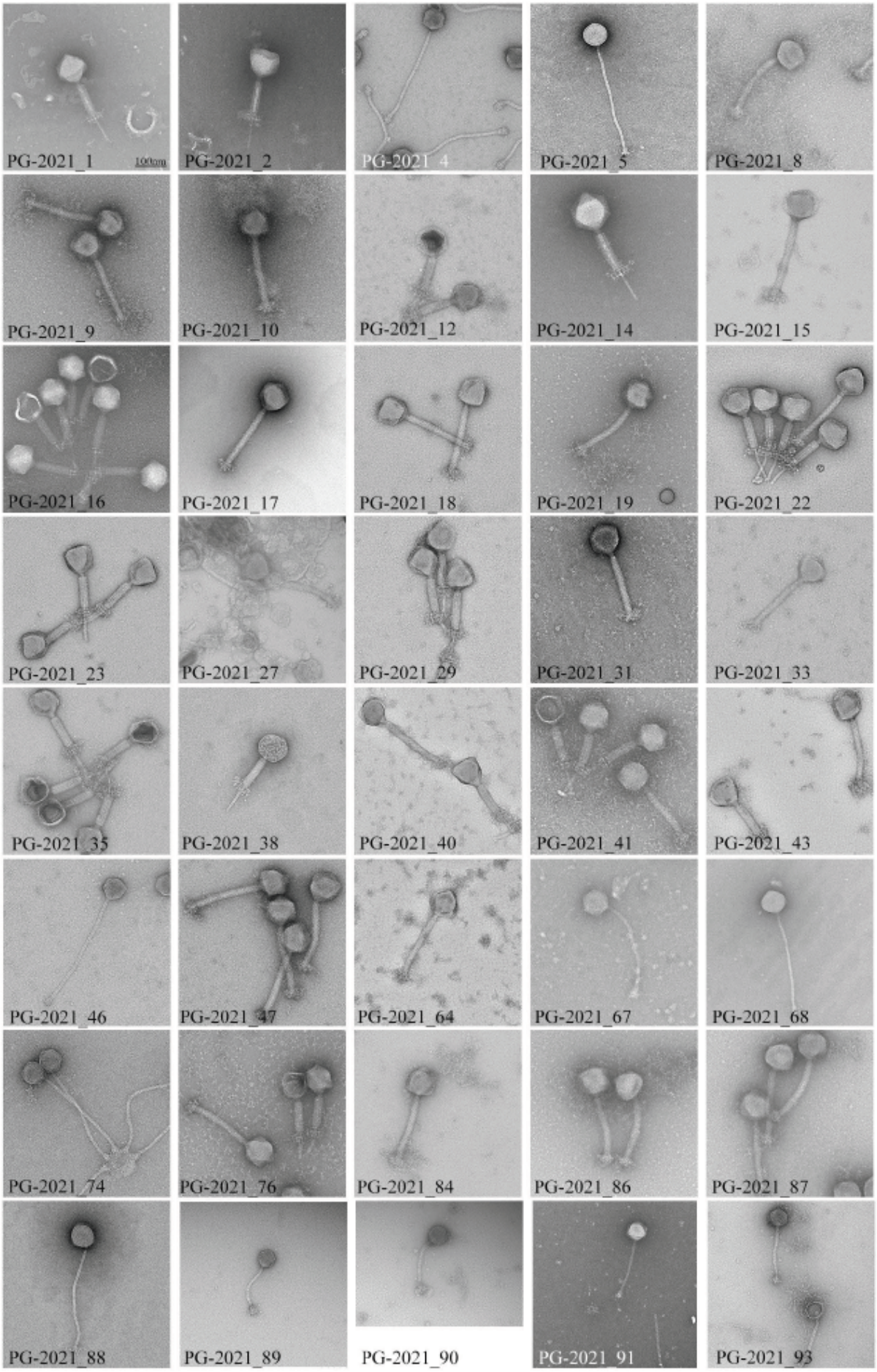

PG-2021_1
100nm
PG-2021_2
PG-2021_4
PG-2021_5
PG-2021_8
PG-2021_9
PG-2021_10
PG-2021_12
PG-2021_14
PG-2021_15
PG-2021_16
PG-2021_17
PG-2021_18
PG-2021_19
PG-2021_22
PG-2021_23
PG-2021_27
PG-2021_29
PG-2021_31
PG-2021_33
PG-2021_35
PG-2021_38
PG-2021_40
PG-2021_41
PG-2021_43
PG-2021_46
PG-2021_47
PG-2021_64
PG-2021_67
PG-2021_68
PG-2021_74
PG-2021_76
PG-2021_84
PG-2021_86
PG-2021_87
PG-2021_88
PG-2021_89
PG-2021_90
PG-2021_91
PG-2021_93

Fig 5

Electron micrographs of the sequenced staphylococcal phages. Phages were isolated from the wastewater treatment plant inlet, outlet, or by induction of bacterial lysogens. All pictures are adjusted according to the displayed scale-bar on the top-left corner.

Figure 6

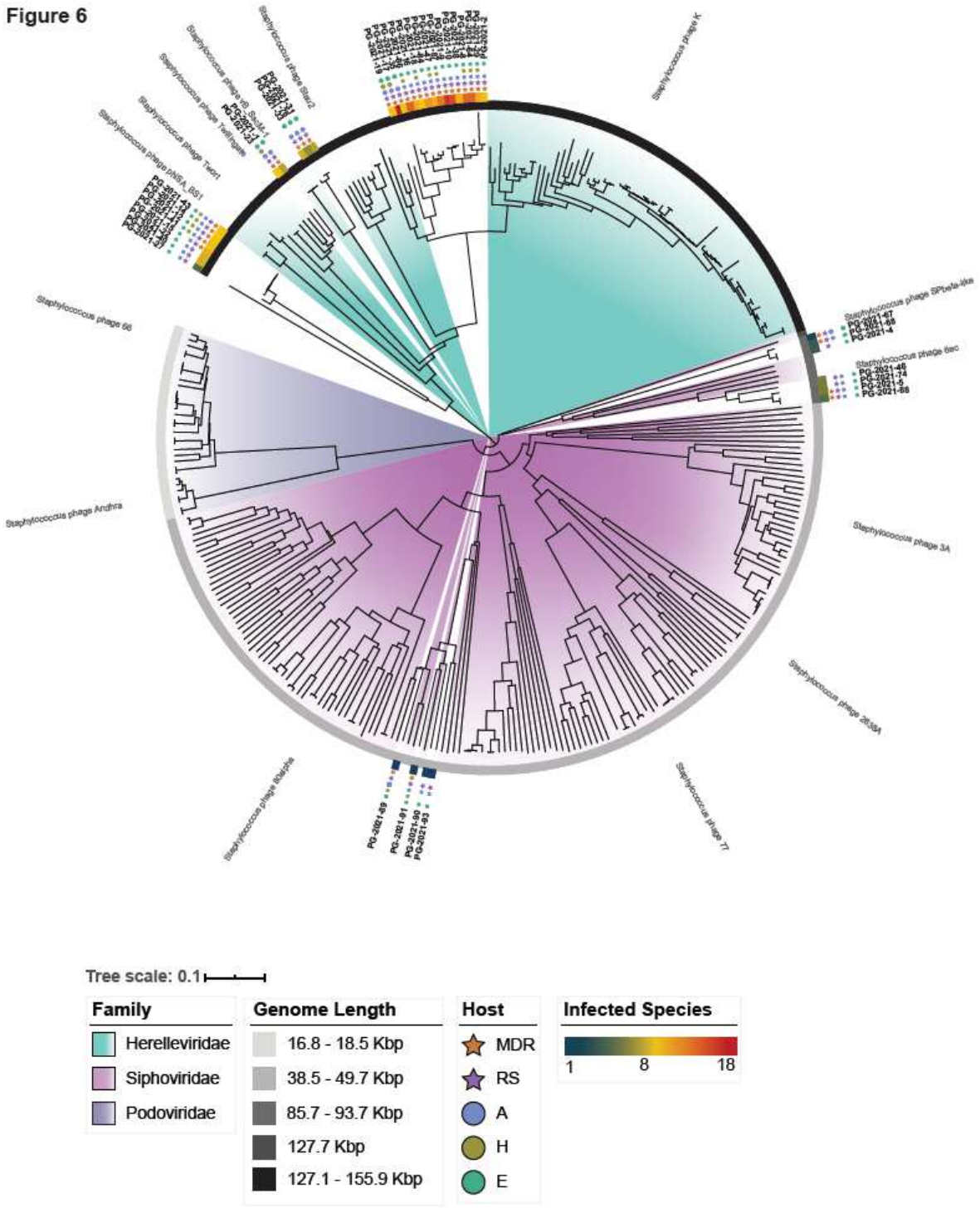
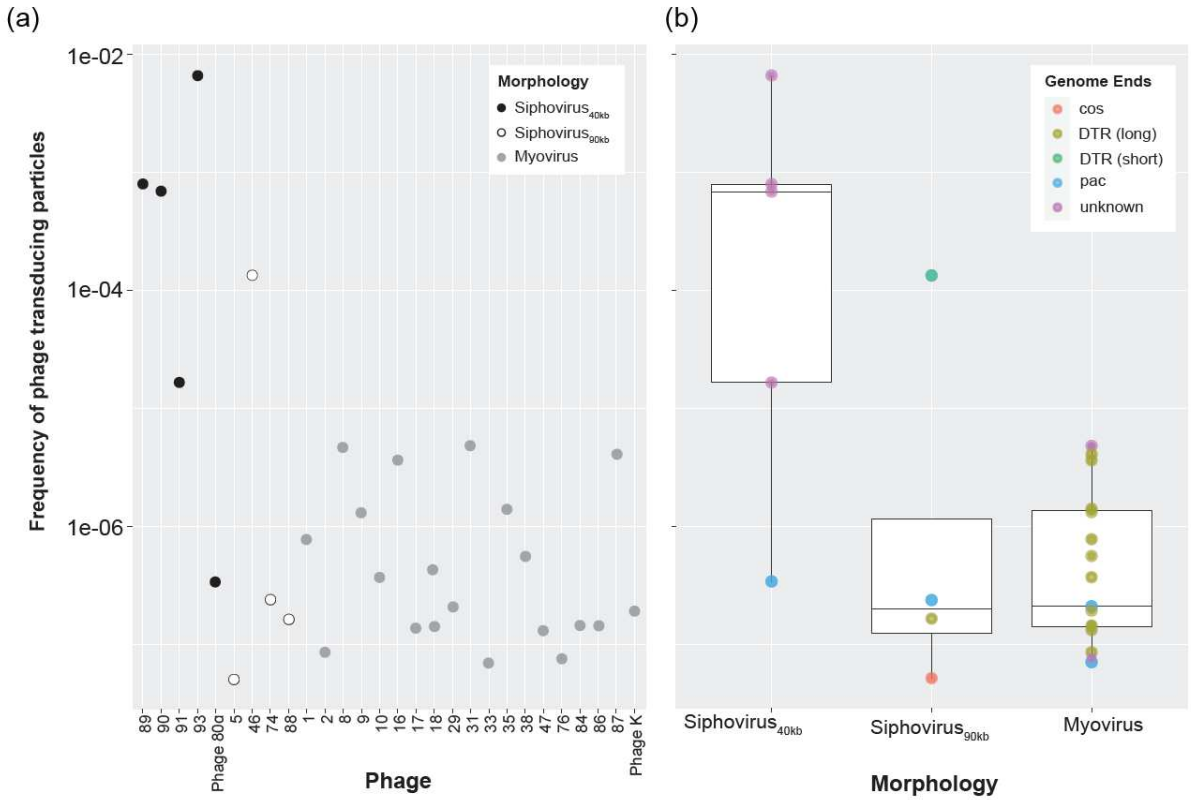


Fig 6

Phylogenomic tree of staphylococcal phages. All published staphylococcal phages (Supplementary Table 15) are displayed together with the here isolated and sequenced CoNS-infecting viruses. For each phage genus, a representative phage is indicated. Phages from our collection are represented in bold and their corresponding host range is represented as follows: the number of infected species is indicated using a continued color scale; isolation origin, and antimicrobial resistant phenotype of infected hosts are represented using colored circles and

798 stars, respectively. Phages infect A: hosts isolated from animals, H: hosts isolated from
799 humans, E: hosts isolated from the environment. MDR: host is multidrug resistant. RS: phage
800 infects hosts with antimicrobial resistant and susceptible phenotypes.

801



802

803 **Fig 7**

804 **Frequency of transducing particles for diverse staphylococcal phages.** (a) Estimated
805 frequency of transducing particles for each respective phage and corresponding phage
806 morphology. Phages are abbreviated with their final unique numerical identifier (PG-2021_*).
807 (b) Mean frequencies of transducing particles for each phage morphology. Phage termini were
808 detected using PhageTerm[96] and are illustrated using colors. DTR: direct terminal repeats.

809

Supplementary Fig 1

Matrix representation of the modular and nested network structure. The matrix is composed of 60 phage permissive staphylococcal strains from 27 species and 94 phages. The rows represent bacteria, and columns represent phages. Grey cells illustrate reported infections. (a) Illustration of the modular sorting. Infections within modules are represented in color. The modularity level (Q), estimated with the lpbrim package in R, is indicated on the bottom line. (b) Illustration of the nestedness sorting. The matrix is now arranged to maximize nestedness. The nestedness, estimated with the NODF function, is indicated in the lower right corner. Both algorithms are described and explained in[14].

Supplementary Fig 2

Host susceptibility towards phage infection. Depicted are the number of phage resistant and permissive hosts that are clustered according to their (a) isolation origin and (b) antimicrobial resistant phenotype.

Supplementary Fig 3

Illustration of the host ranges collapsed on the species level for all isolated phages. Phages on the x-axis are sorted from narrow (left) to broad host range (right). Species on the y-axis are sorted after phylogenetic relationship in species groups[18]. A phage host range is depicted as a column, where infection of a staphylococcal species is illustrated using circles. For each respective species, the area of the circle is scaled according to the number of strains a phage can replicate on (scale: 1-15). The total number of strains challenged per species is depicted in the bar-chart on the right. Host ranges on this host array are colored as follows: Phage with species tendency (≥ 50 % of all infections on a single species) in turquoise; phages with no species tendency in dark blue; polyvalent phage K in violet. Phages are abbreviated with their final unique numerical identifier (PG-2021_*).

Supplementary Fig 4

The number of shared phages between hosts in the bipartite network projection. With our natural phage community being present, two bacterial hosts share between zero and 58 phages. The relative frequency indicates how many bacterial hosts share the respective number of phages (bin width two). On average, each host pair is connected by 4.2 ± 7.9 (n=1770) different phages (mean ± sd in red).

Supplementary Fig 5

Phage susceptibility of strains classified after drug resistance phenotypes. For each strain, the number of total phages infecting this host was determined and depicted on the y-axis. On the x-axis, strains are classified according to their antimicrobial resistance phenotype.

The average number of phages infecting a host phenotype is illustrated in red (mean±sd). There is no significant difference in the infection of antimicrobial susceptible or resistant hosts (Two-sided Wilcoxon rank sum test with continuity correction, W = 568.5, n = 60, p-value = 0.068).

Supplementary Fig 6

Phages connecting hosts of different antimicrobial resistant phenotypes and epidemiologic backgrounds. (a) Phages infecting either exclusively drug resistant, or drug resistant and susceptible bacteria are represented as rows. The number of phages connecting hosts from the environmental ecosystem, both the environmental and veterinary ecosystem, or all three ecosystems are represented as columns. (b) The average number of shared phages (mean±sd) between hosts within or across an ecosystem is depicted. Staphylococcal hosts are classified according to their isolation origin into environmental, veterinary, or human associated strains.

Supplementary Fig 7

Mean neighbor count of a phage permissive host in the bipartite network projection. Phage permissive hosts (60) were categorized after their isolation origin, and the number of direct neighbors connected through phages was counted. Neighbors themselves were subdivided according to their isolation origin. On average, hosts isolated from animals revealed to have 33.5 ± 14.5 (n = 53) neighbors, environmental hosts 39.9 ± 11.6 (n = 23), hosts isolated from the human biome 24 ± 9 (n = 40), and hosts of unknown isolation origin 37.8 ± 12.6 (n = 7) neighbors.

References

1 Roux, S. *et al.* Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* **537**, 689-693, doi:10.1038/nature19366 (2016).

2 Kauffman, K. M. *et al.* A major lineage of non-tailed dsDNA viruses as unrecognized killers of marine bacteria. *Nature* **554**, 118-122, doi:10.1038/nature25474 (2018).

3 Touchon, M., Moura de Sousa, J. A. & Rocha, E. P. Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer. *Curr. Opin. Microbiol.* **38**, 66-73, doi:10.1016/j.mib.2017.04.010 (2017).

4 Chiang, Y. N., Penades, J. R. & Chen, J. Genetic transduction by phages and chromosomal islands: The new and noncanonical. *PLoS Pathog.* **15**, e1007878, doi:10.1371/journal.ppat.1007878 (2019).

5 Penades, J. R., Chen, J., Quiles-Puchalt, N., Carpena, N. & Novick, R. P. Bacteriophage-mediated spread of bacterial virulence genes. *Curr. Opin. Microbiol.* **23**, 171-178, doi:10.1016/j.mib.2014.11.019 (2015).

6 de Jonge, P. A., Nobrega, F. L., Brouns, S. J. J. & Dutilh, B. E. Molecular and Evolutionary Determinants of Bacteriophage Host Range. *Trends Microbiol.* **27**, 51-63, doi:10.1016/j.tim.2018.08.006 (2019).

7 Mahony, J., Casey, E. & van Sinderen, D. The Impact and Applications of Phages in the Food Industry and Agriculture. *Viruses* **12**, doi:10.3390/v12020210 (2020).

8 Cisek, A. A., Dabrowska, I., Gregorczyk, K. P. & Wyzewski, Z. Phage Therapy in Bacterial Infections Treatment: One Hundred Years After the Discovery of Bacteriophages. *Curr. Microbiol.* **74**, 277-283, doi:10.1007/s00284-016-1166-x (2017).

9 Mohan Raj, J. R. & Karunasagar, I. Phages amid antimicrobial resistance. *Crit. Rev. Microbiol.* **45**, 701-711, doi:10.1080/1040841X.2019.1691973 (2019).

10 Bernheim, A. & Sorek, R. The pan-immune system of bacteria: antiviral defence as a community resource. *Nat. Rev. Microbiol.* **18**, 113-119, doi:10.1038/s41579-019-0278-2 (2020).

11 Munson-McGee, J. H. *et al.* A virus or more in (nearly) every cell: ubiquitous networks of virus-host interactions in extreme environments. *Isme J* **12**, 1706-1714, doi:10.1038/s41396-018-0071-7 (2018).

12 Dion, M. B., Oechslin, F. & Moineau, S. Phage diversity, genomics and phylogeny. *Nat. Rev. Microbiol.* **18**, 125-138, doi:10.1038/s41579-019-0311-5 (2020).

13 Flores, C. O., Valverde, S. & Weitz, J. S. Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages. *Isme J* **7**, 520-532, doi:10.1038/ismej.2012.135 (2013).

14 Weitz, J. S. *et al.* Phage-bacteria infection networks. *Trends Microbiol.* **21**, 82-91, doi:10.1016/j.tim.2012.11.003 (2013).

15 Flores, C. O., Meyer, J. R., Valverde, S., Farr, L. & Weitz, J. S. Statistical structure of host-phage interactions. *Proc. Natl. Acad. Sci. U. S. A.* **108**, E288-297, doi:10.1073/pnas.1101595108 (2011).

16 Haaber, J., Penades, J. R. & Ingmer, H. Transfer of Antibiotic Resistance in Staphylococcus aureus. *Trends Microbiol.* **25**, 893-905, doi:10.1016/j.tim.2017.05.011 (2017).

17 Becker, K., Heilmann, C. & Peters, G. Coagulase-negative staphylococci. *Clin. Microbiol. Rev.* **27**, 870-926, doi:10.1128/CMR.00109-13 (2014).

18 Lamers, R. P. *et al.* Phylogenetic relationships among Staphylococcus species and refinement of cluster groups based on multilocus data. *BMC Evol. Biol.* **12**, 171, doi:10.1186/1471-2148-12-171 (2012).

19 Oliveira, H. *et al.* Staphylococci phages display vast genomic diversity and evolutionary relationships. *BMC Genomics* **20**, 357, doi:10.1186/s12864-019-5647-8 (2019).

20 Moller, A. G., Lindsay, J. A. & Read, T. D. Determinants of Phage Host Range in Staphylococcus Species. *Appl. Environ. Microbiol.* **85**, doi:10.1128/AEM (2019).

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

22 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, doi:10.1128/mBio.00869-14 (2014).

23 Xia, G., Kohler, T. & Peschel, A. The wall teichoic acid and lipoteichoic acid polymers of Staphylococcus aureus. *Int. J. Med. Microbiol.* **300**, 148-154, doi:10.1016/j.ijmm.2009.10.001 (2010).

24 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).

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

26 Uchiyama, J. *et al.* Intragenus generalized transduction in Staphylococcus spp. by a novel giant phage. *Isme J* **8**, 1949-1952, doi:10.1038/ismej.2014.29 (2014).

27 Melo, L. D. R., Brandao, A., Akturk, E., Santos, S. B. & Azeredo, J. Characterization of a New Staphylococcus aureus Kayvirus Harboring a Lysin Active against Biofilms. *Viruses* **10**, doi:10.3390/v10040182 (2018).

28 Oduor, J. M. O., Kadija, E., Nyachieo, A., Mureithi, M. W. & Skurnik, M. Bioprospecting Staphylococcus Phages with Therapeutic and Bio-Control Potential. *Viruses* **12**, doi:10.3390/v12020133 (2020).

29 Deghorain, M. *et al.* Characterization of novel phages isolated in coagulase-negative staphylococci reveals evolutionary relationships with Staphylococcus aureus phages. *J. Bacteriol.* **194**, 5829-5839, doi:10.1128/JB.01085-12 (2012).

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

31 Lobocka, M. *et al.* Genomics of staphylococcal Twort-like phages--potential therapeutics of the post-antibiotic era. *Adv. Virus Res.* **83**, 143-216, doi:10.1016/B978-0-12-394438-2.00005-0 (2012).

32 Hyman, P. & Abedon, S. T. in *Bacteriophage Host Range and Bacterial Resistance Advances in Applied Microbiology* 217-248 (2010).

33 Bailly-Bechet, M., Vergassola, M. & Rocha, E. Causes for the intriguing presence of tRNAs in phages. *Genome Res.* **17**, 1486-1495, doi:10.1101/gr.6649807 (2007).

34 Gómez-Sanz, E., Haro-Moreno, J. M., Jensen, S. O., Roda-García, J. J. & López-Pérez, M. *Staphylococcus sciuri* C2865 from a distinct subspecies cluster as reservoir of the novel transferable trimethoprim resistance gene, *dfrE*, and adaptation driving mobile elements. *bioRxiv*, 2020.2009.2030.320143, doi:10.1101/2020.09.30.320143 (2020).

35 Colomer-Lluch, M. *et al.* Antibiotic resistance genes in bacterial and bacteriophage fractions of Tunisian and Spanish wastewaters as markers to compare the antibiotic

resistance patterns in each population. *Environ. Int.* **73**, 167-175, doi:10.1016/j.envint.2014.07.003 (2014).

36 Novick, R. P., Edelman, I. & Lofdahl, S. Small Staphylococcus aureus Plasmids are Transduced as Linear Multimers that are Formed and Resolved by Replicative Processes. *J. Mol. Biol.* **192**, 209-220 (1986).

37 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).

38 Varga, M. *et al.* Efficient transfer of antibiotic resistance plasmids by transduction within methicillin-resistant Staphylococcus aureus USA300 clone. *FEMS Microbiol. Lett.* **332**, 146-152, doi:10.1111/j.1574-6968.2012.02589.x (2012).

39 Maslanova, I. *et al.* Bacteriophages of Staphylococcus aureus efficiently package various bacterial genes and mobile genetic elements including SCCmec with different frequencies. *Environ. Microbiol. Rep.* **5**, 66-73, doi:10.1111/j.1758-2229.2012.00378.x (2013).

40 Rodriguez-Rubio, L. *et al.* Extensive antimicrobial resistance mobilization via multicopy plasmid encapsidation mediated by temperate phages. *J. Antimicrob. Chemother.*, doi:10.1093/jac/dkaa311 (2020).

41 Chen, J. *et al.* Genome hypermobility by lateral transduction. *Science* **362**, 207-212, doi:10.1126/science.aat5867 (2018).

42 Zeman, M. *et al.* New Genus Fibralongavirus in Siphoviridae Phages of Staphylococcus pseudintermedius. *Viruses* **11**, doi:10.3390/v11121143 (2019).

43 Hsieh, S. E., Lo, H. H., Chen, S. T., Lee, M. C. & Tseng, Y. H. Wide host range and strong lytic activity of Staphylococcus aureus lytic phage Stau2. *Appl. Environ. Microbiol.* **77**, 756-761, doi:10.1128/AEM.01848-10 (2011).

44 Synnott, A. J. *et al.* Isolation from sewage influent and characterization of novel Staphylococcus aureus bacteriophages with wide host ranges and potent lytic capabilities. *Appl. Environ. Microbiol.* **75**, 4483-4490, doi:10.1128/AEM.02641-08 (2009).

45 Kvachadze, L. *et al.* Evaluation of lytic activity of staphylococcal bacteriophage Sb-1 against freshly isolated clinical pathogens. *Microb. Biotechnol.* **4**, 643-650, doi:10.1111/j.1751-7915.2011.00259.x (2011).

46 Gutierrez, D. *et al.* Phage sensitivity and prophage carriage in Staphylococcus aureus isolated from foods in Spain and New Zealand. *Int. J. Food Microbiol.* **230**, 16-20, doi:10.1016/j.ijfoodmicro.2016.04.019 (2016).

47 Gutierrez, D., Martınez, B., Rodrıguez, A. & Garcıa, P. Isolation and Characterization of Bacteriophages Infecting Staphylococcus epidermidis. *Curr. Microbiol.* **61**, 601–608, doi:10.1007/s00284-010-9659-5 (2010).

48 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).

49 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).

50 Zeman, M. *et al.* Staphylococcus sciuri bacteriophages double-convert for staphylokinase and phospholipase, mediate interspecies plasmid transduction, and package mecA gene. *Sci. Rep.* **7**, 46319, doi:10.1038/srep46319 (2017).

51 Chen, J. & Novick, R. P. Phage-Mediated Intergeneric Transfer of Toxin Genes. *Science* **323** (2009).

52 Nair, D. *et al.* Whole-genome sequencing of Staphylococcus aureus strain RN4220, a key laboratory strain used in virulence research, identifies mutations that affect not only virulence factors but also the fitness of the strain. *J. Bacteriol.* **193**, 2332-2335, doi:10.1128/JB.00027-11 (2011).

53 Aspiroz, C. *et al.* Skin Lesion Caused by ST398 and ST1 MRSA, Spain1. *Emerg. Infect. Dis.* **16**, 156-157, doi:10.3201/eid1601.091420 (2010).

54 Unlu, G., Nielsen, B. & Ionita, C. Production of Antilisterial Bacteriocins from Lactic Acid Bacteria in Dairy-Based Media: A Comparative Study. *Probiotics Antimicrob Proteins* **7**, 259-274, doi:10.1007/s12602-015-9200-z (2015).

55 Gandolfi-Decristophoris, P., Regula, G., Petrini, O., Zinsstag, J. & Schelling, E. Prevalence and risk factors for carriage of multi-drug resistantStaphylococciin healthy cats and dogs. *J. Vet. Sci.* **14**, 449, doi:10.4142/jvs.2013.14.4.449 (2012).

56 Schwendener, S., Cotting, K. & Perreten, V. Novel methicillin resistance gene mecD in clinical Macrococcus caseolyticus strains from bovine and canine sources. *Sci. Rep.* **7**, 43797, doi:10.1038/srep43797 (2017).

57 Wipf, J. R., Schwendener, S. & Perreten, V. The novel macrolide-Lincosamide-Streptogramin B resistance gene erm(44) is associated with a prophage in Staphylococcus xylosus. *Antimicrob. Agents Chemother.* **58**, 6133-6138, doi:10.1128/AAC.02949-14 (2014).

58 Wipf, J. R. K. *et al.* New Macrolide-Lincosamide-Streptogramin B Resistance Gene erm(48) on the Novel Plasmid pJW2311 in Staphylococcus xylosus. *Antimicrob. Agents Chemother.* **61** (2017).

59 Wipf, J. R., Schwendener, S., Nielsen, J. B., Westh, H. & Perreten, V. The new macrolide-lincosamide-streptogramin B resistance gene erm(45) is located within a genomic island in Staphylococcus fleurettii. *Antimicrob. Agents Chemother.* **59**, 3578-3581, doi:10.1128/AAC.00369-15 (2015).

60 Ben Slama, K. *et al.* Nasal carriage of Staphylococcus aureus in healthy humans with different levels of contact with animals in Tunisia: genetic lineages, methicillin resistance, and virulence factors. *Eur. J. Clin. Microbiol. Infect. Dis.* **30**, 499-508, doi:10.1007/s10096-010-1109-6 (2011).

61 Chah, K. F. *et al.* Methicillin-resistant coagulase-negative staphylococci from healthy dogs in Nsukka, Nigeria. *Braz. J. Microbiol.* **45**, 215-220 (2014).

62 Cotting, K. *et al.* Macrococcus canis and M. caseolyticus in dogs: occurrence, genetic diversity and antibiotic resistance. *Vet. Dermatol.* **28**, 559-e133, doi:10.1111/vde.12474 (2017).

63 Verbree, C. T. *et al.* Identification of Peptidoglycan Hydrolase Constructs with Synergistic Staphylolytic Activity in Cow's Milk. *Appl. Environ. Microbiol.* **83**, doi:10.1128/AEM.03445-16 (2017).

64 Gomez-Sanz, E., Torres, C., Lozano, C. & Zarazaga, M. High diversity of Staphylococcus aureus and Staphylococcus pseudintermedius lineages and toxigenic traits in healthy pet-owning household members. Underestimating normal household contact? *Comp. Immunol. Microbiol. Infect. Dis.* **36**, 83-94, doi:10.1016/j.cimid.2012.10.001 (2013).

65 Baba, T., Bae, T., Schneewind, O., Takeuchi, F. & Hiramatsu, K. Genome sequence of Staphylococcus aureus strain Newman and comparative analysis of staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands. *J. Bacteriol.* **190**, 300-310, doi:10.1128/JB.01000-07 (2008).

66 Benito, D. *et al.* Genetic lineages and antimicrobial resistance genotypes in Staphylococcus aureus from children with atopic dermatitis: detection of clonal complexes CC1, CC97 and CC398. *J. Chemother.* **28**, 359-366, doi:10.1179/1973947815Y.0000000044 (2016).

67 Winstel, V., Kuhner, P., Rohde, H. & Peschel, A. Genetic engineering of untransformable coagulase-negative staphylococcal pathogens. *Nat. Protoc.* **11**, 949-959, doi:10.1038/nprot.2016.058 (2016).

68 Frey, Y., Rodriguez, J. P., Thomann, A., Schwendener, S. & Perreten, V. Genetic characterization of antimicrobial resistance in coagulase-negative staphylococci from bovine mastitis milk. *J. Dairy Sci.* **96**, 2247-2257, doi:10.3168/jds.2012-6091 (2013).

69 Gomez-Sanz, E., Schwendener, S., Thomann, A., Gobeli Brawand, S. & Perreten, V. First Staphylococcal Cassette Chromosome mec Containing a mecB-Carrying Gene Complex Independent of Transposon Tn6045 in a Macrococcus canis Isolate from a Canine Infection. *Antimicrob. Agents Chemother.* **59**, 4577-4583, doi:10.1128/AAC.05064-14 (2015).

70 Gomez-Sanz, E. *et al.* First detection of methicillin-resistant Staphylococcus aureus ST398 and Staphylococcus pseudintermedius ST68 from hospitalized equines in Spain. *Zoonoses Public Health* **61**, 192-201, doi:10.1111/zph.12059 (2014).

71 Gomez, P. *et al.* Diversity of species and antimicrobial resistance determinants of staphylococci in superficial waters in Spain. *FEMS Microbiol. Ecol.* **93**, doi:10.1093/femsec/fiw208 (2017).

72 Gomez-Sanz, E. *et al.* Detection, molecular characterization, and clonal diversity of methicillin-resistant Staphylococcus aureus CC398 and CC97 in Spanish slaughter pigs of different age groups. *Foodborne Pathog. Dis.* **7**, 1269-1277, doi:10.1089/fpd.2010.0610 (2010).

73 Gomez, P. *et al.* Detection of MRSA ST3061-t843-mecC and ST398-t011-mecA in white stork nestlings exposed to human residues. *J. Antimicrob. Chemother.* **71**, 53-57, doi:10.1093/jac/dkv314 (2016).

74 Benito, D., Lozano, C., Gomez-Sanz, E., Zarazaga, M. & Torres, C. Detection of methicillin-susceptible Staphylococcus aureus ST398 and ST133 strains in gut microbiota of healthy humans in Spain. *Microb. Ecol.* **66**, 105-111, doi:10.1007/s00248-013-0240-1 (2013).

75 Lozano, C. *et al.* Detection of methicillin-resistant Staphylococcus aureus ST398 in food samples of animal origin in Spain. *J. Antimicrob. Chemother.* **64**, 1325-1326, doi:10.1093/jac/dkp378 (2009).

76 Gomez, P. *et al.* Detection of methicillin-resistant Staphylococcus aureus (MRSA) carrying the mecC gene in wild small mammals in Spain. *J. Antimicrob. Chemother.* **69**, 2061-2064, doi:10.1093/jac/dku100 (2014).

77 Gomez-Sanz, E., Torres, C., Lozano, C., Saenz, Y. & Zarazaga, M. Detection and characterization of methicillin-resistant Staphylococcus pseudintermedius in healthy dogs in La Rioja, Spain. *Comp. Immunol. Microbiol. Infect. Dis.* **34**, 447-453, doi:10.1016/j.cimid.2011.08.002 (2011).

78 Diep, B. A. *et al.* Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant Staphylococcus aureus. *The Lancet* **367**, 731-739, doi:10.1016/s0140-6736(06)68231-7 (2006).

79 Gomez-Sanz, E., Ceballos, S., Ruiz-Ripa, L., Zarazaga, M. & Torres, C. Clonally Diverse Methicillin and Multidrug Resistant Coagulase Negative Staphylococci Are Ubiquitous and Pose Transfer Ability Between Pets and Their Owners. *Front. Microbiol.* **10**, 485, doi:10.3389/fmicb.2019.00485 (2019).

80 Gomez-Sanz, E., Torres, C., Ceballos, S., Lozano, C. & Zarazaga, M. Clonal dynamics of nasal Staphylococcus aureus and Staphylococcus pseudintermedius in dog-owning household members. Detection of MSSA ST(398). *PLoS One* **8**, e69337, doi:10.1371/journal.pone.0069337 (2013).

81 Benito, D. *et al.* Characterization of Staphylococcus aureus strains isolated from faeces of healthy neonates and potential mother-to-infant microbial transmission through breastfeeding. *FEMS Microbiol. Ecol.* **91**, doi:10.1093/femsec/fiv007 (2015).

82 Gomez, P. *et al.* Characterization of staphylococci in urban wastewater treatment plants in Spain, with detection of methicillin resistant Staphylococcus aureus ST398. *Environ. Pollut.* **212**, 71-76, doi:10.1016/j.envpol.2016.01.038 (2016).

83 Lozano, C. *et al.* [Characterization of methicillin- and linezolid-resistant Staphylococcus epidermidis and S. haemolyticus strains in a Spanish hospital]. *Enferm. Infecc. Microbiol. Clin.* **31**, 136-141, doi:10.1016/j.eimc.2012.08.006 (2013).

84 Ugwu, C. C., Gomez-Sanz, E., Agbo, I. C., Torres, C. & Chah, K. F. Characterization of mannitol-fermenting methicillin-resistant staphylococci isolated from pigs in Nigeria. *Braz. J. Microbiol.* **46**, 885-892, doi:10.1590/S1517-838246320140644 (2015).

85 Lozano, C. *et al.* Characterization of a cfr-positive methicillin-resistant Staphylococcus epidermidis strain of the lineage ST22 implicated in a life-threatening human infection. *Diagn. Microbiol. Infect. Dis.* **73**, 380-382, doi:10.1016/j.diagmicrobio.2012.04.013 (2012).

86 Clokie, M. R. J. & Kropinski, A. *Bacteriophages*. Vol. 1 (Humana Press, 2009).

87 Bae, T., Baba, T., Hiramatsu, K. & Schneewind, O. Prophages of Staphylococcus aureus Newman and their contribution to virulence. *Mol. Microbiol.* **62**, 1035-1047, doi:10.1111/j.1365-2958.2006.05441.x (2006).

88 Coman, M. M. *et al.* In vitro evaluation of antimicrobial activity of Lactobacillus rhamnosus IMC 501((R)) , Lactobacillus paracasei IMC 502((R)) and SYNBIO((R)) against pathogens. *J. Appl. Microbiol.* **117**, 518-527, doi:10.1111/jam.12544 (2014).

89 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).

90 Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455-477, doi:10.1089/cmb.2012.0021 (2012).

91 Chin, C. S. *et al.* Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* **10**, 563-569, doi:10.1038/nmeth.2474 (2013).

92 McNair, K., Zhou, C., Dinsdale, E. A., Souza, B. & Edwards, R. A. PHANOTATE: a novel approach to gene identification in phage genomes. *Bioinformatics* **35**, 4537-4542, doi:10.1093/bioinformatics/btz265 (2019).

93 Ecale Zhou, C. L. *et al.* multiPhATE: bioinformatics pipeline for functional annotation of phage isolates. *Bioinformatics* **35**, 4402-4404, doi:10.1093/bioinformatics/btz258 (2019).

94 Grazziotin, A. L., Koonin, E. V. & Kristensen, D. M. Prokaryotic Virus Orthologous Groups (pVOGs): a resource for comparative genomics and protein family annotation. *Nucleic Acids Res.* **45**, D491-D498, doi:10.1093/nar/gkw975 (2017).

95 Chan, P. P. & Lowe, T. M. tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. *Methods Mol. Biol.* **1962**, 1-14, doi:10.1007/978-1-4939-9173-0_1 (2019).

96 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).

97 Richter, M. & Rossello-Mora, R. Shifting the genomic gold standard for the prokaryotic species definition. *PNAS* **106**, 19126–19131 (2009).

98 Tange, O. *GNU Parallel*, 2018).

99 Coutinho, F. H., Edwards, R. A. & Rodriguez-Valera, F. Charting the diversity of uncultured viruses of Archaea and Bacteria. *BMC Biol.* **17**, 109, doi:10.1186/s12915-019-0723-8 (2019).

100 Mizuno, C. M., Rodriguez-Valera, F., Kimes, N. E. & Ghai, R. Expanding the marine virosphere using metagenomics. *PLoS Genet.* **9**, e1003987, doi:10.1371/journal.pgen.1003987 (2013).

101 Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, W256-W259, doi:10.1093/nar/gkz239 (2019).

102 Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJournal, Complex Systems 1695* (2006).

103 lpbrim: LP-BRIM Bipartite Modularity (https://CRAN.R-project.org/package=lpbrim, 2015).

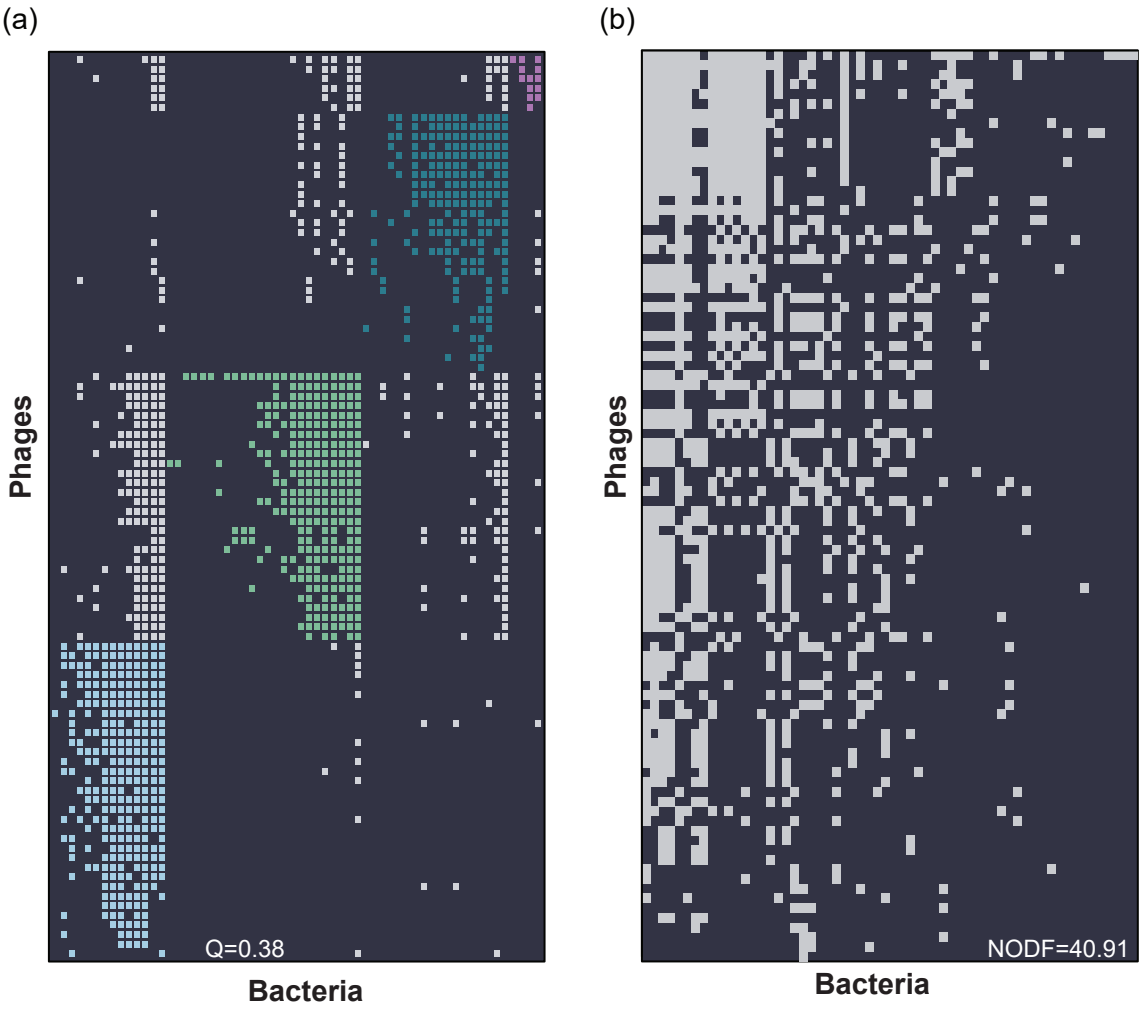
104 vegan: Community Ecology Package (https://CRAN.R-project.org/package=vegan, 2019).

105 Grosser, M. R. & Richardson, A. R. Method for Preparation and Electroporation of S. aureus and S. epidermidis. *Methods Mol. Biol.* **1373**, 51-57, doi:10.1007/7651_2014_183 (2016).

106 Wickham, H. *et al.* Welcome to the Tidyverse. *Journal of Open Source Software* **4**, doi:10.21105/joss.01686 (2019).

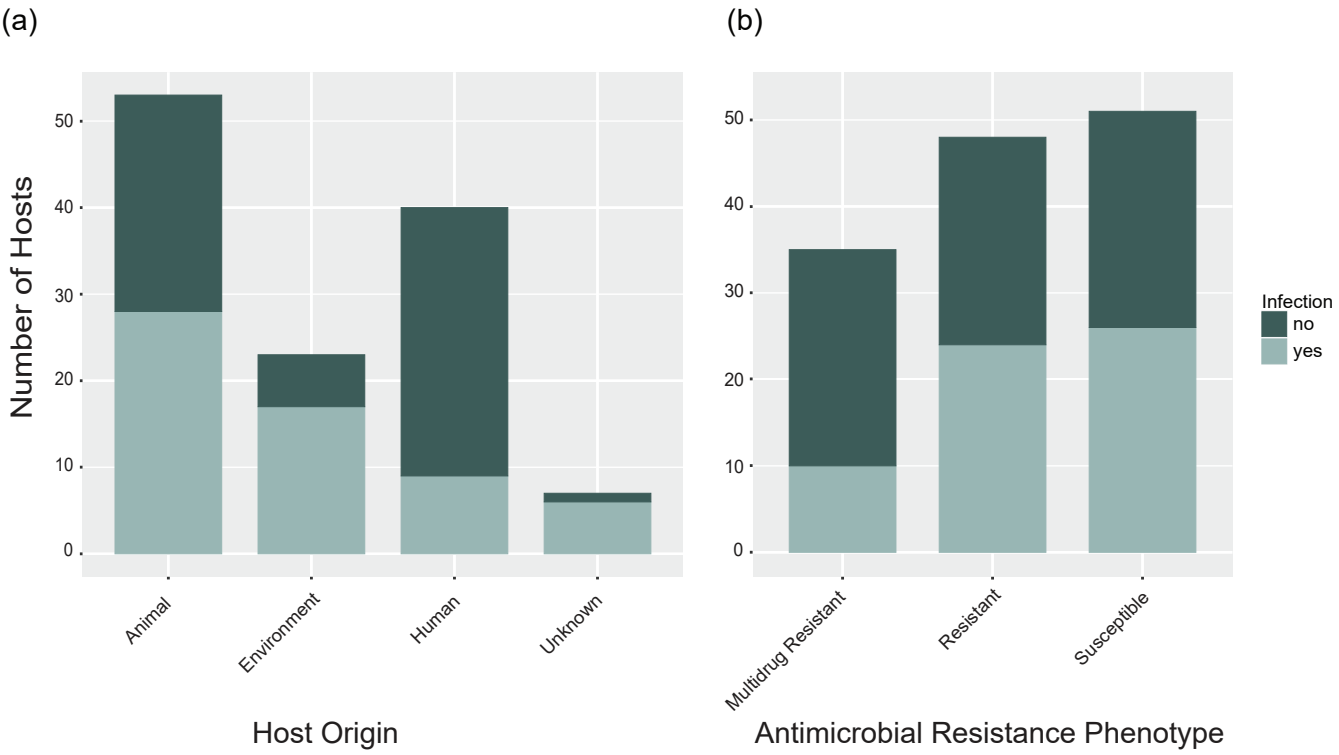
107 Wickham, H. ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York* (2016).

Figure 1



Matrix representation of the modular and nested network structure. The matrix is composed of 60 phage permissive staphylococcal strains from 27 species and 94 phages. The rows represent bacteria, and columns represent phages. Grey cells illustrate reported infections. (a) Illustration of the modular sorting. Infections within modules are represented in color. The modularity level (Q), estimated with the lpbrim package in R, is indicated on the bottom line. (b) Illustration of the nestedness sorting. The matrix is now arranged to maximize nestedness. The nestedness, estimated with the NODF function, is indicated in the lower right corner. Both algorithms are described and explained in[14].

Figure 2



Host susceptibility towards phage infection. Depicted are the number of phage resistant and permissive hosts that are clustered according to their (a) isolation origin and (b) antimicrobial resistant phenotype.

Figure 3

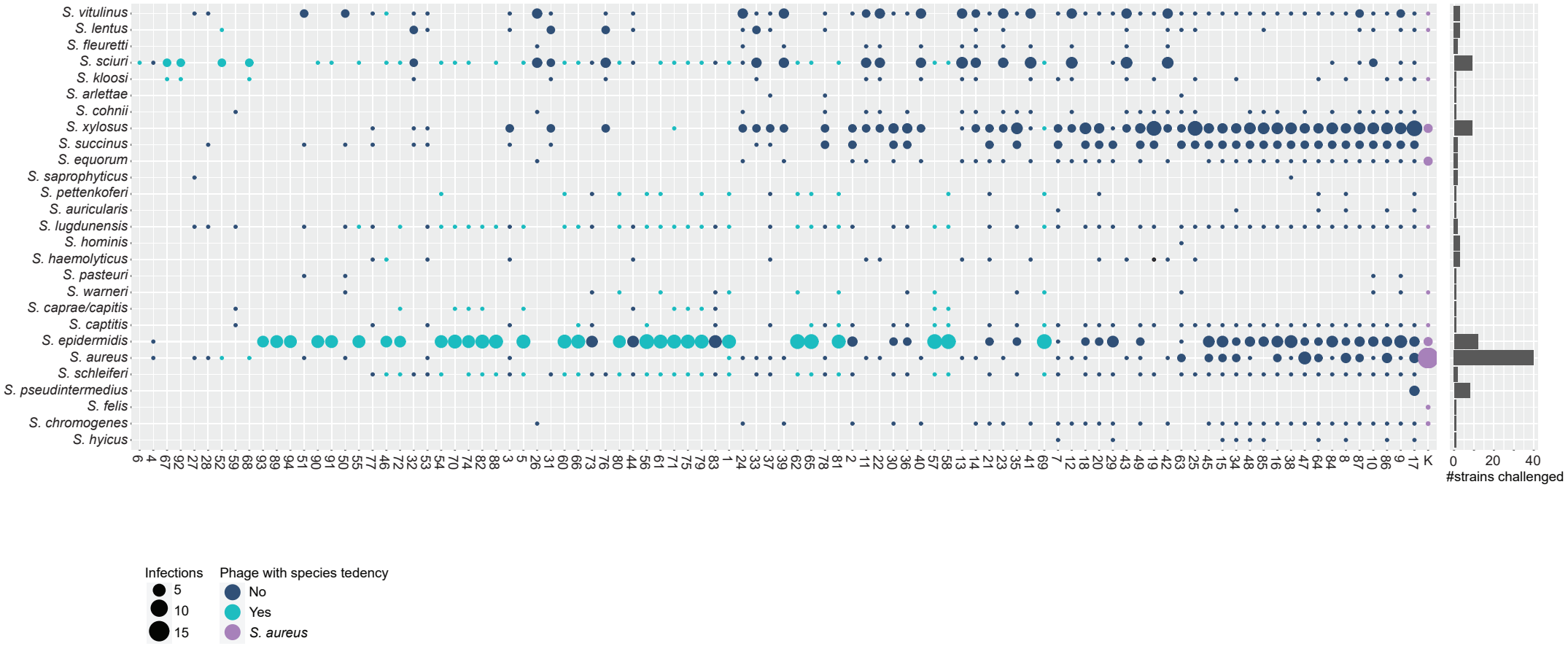
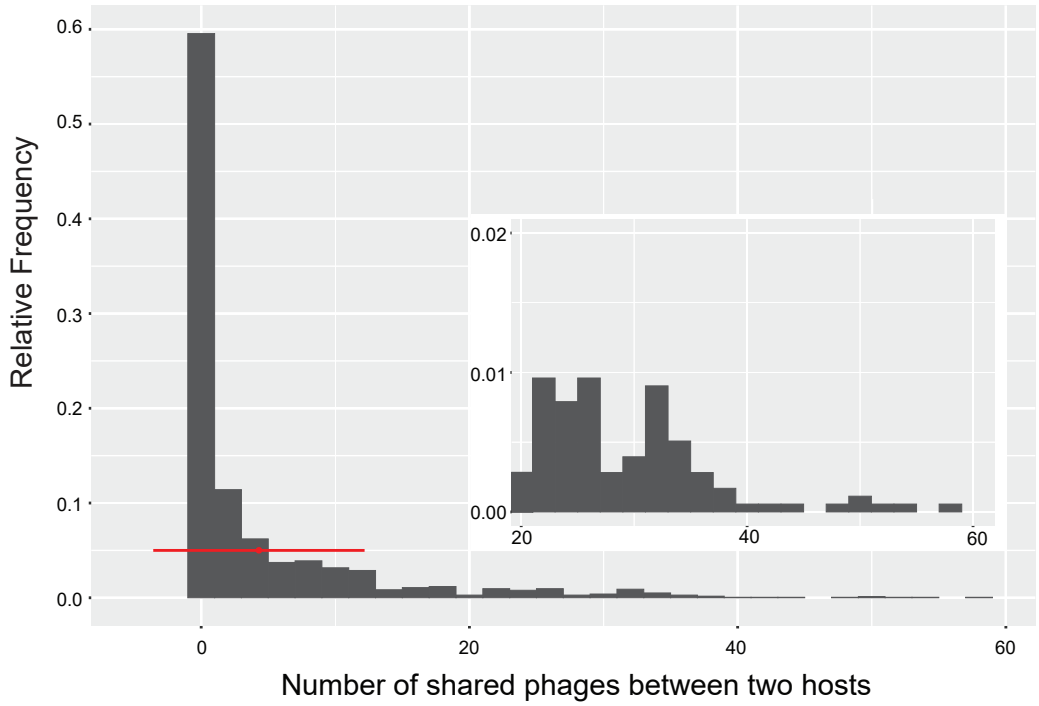


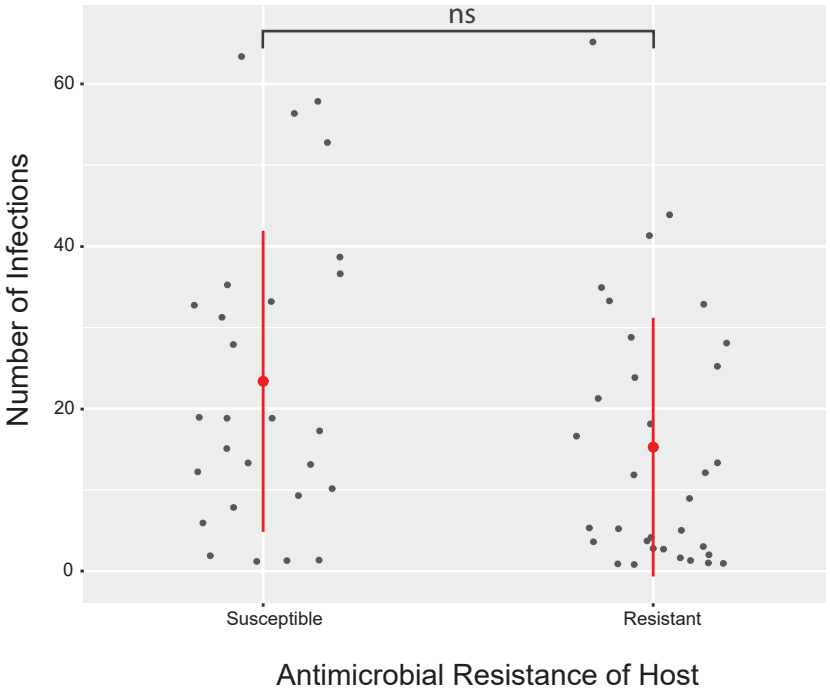
Illustration of the host ranges collapsed on the species level for all isolated phages. Phages on the x-axis are sorted from narrow (left) to broad host range (right). Species on the y-axis are sorted after phylogenetic relationship in species groups[18.] A phage host range is depicted as a column, where infection of a staphylococcal species is illustrated using circles. For each respective species, the area of the circle is scaled according to the number of strains a phage can replicate on (scale: 1-15). The total number of strains challenged per species is depicted in the bar-chart on the right. Host ranges on this host array are colored as follows: Phage with species tendency (≥ 50 % of all infections on a single species) in turquoise; phages with no species tendency in dark blue; polyvalent phage K in violet. Phages are abbreviated with their final unique numerical identifier (PG-2021_*).

Figure 4



The number of shared phages between hosts in the bipartite network projection. With our natural phage community being present, two bacterial hosts share between zero and 58 phages. The relative frequency indicates how many bacterial hosts share the respective number of phages (bin width two). On average, each host pair is connected by 4.2 ± 7.9 (n=1770) different phages (mean ± sd in red).

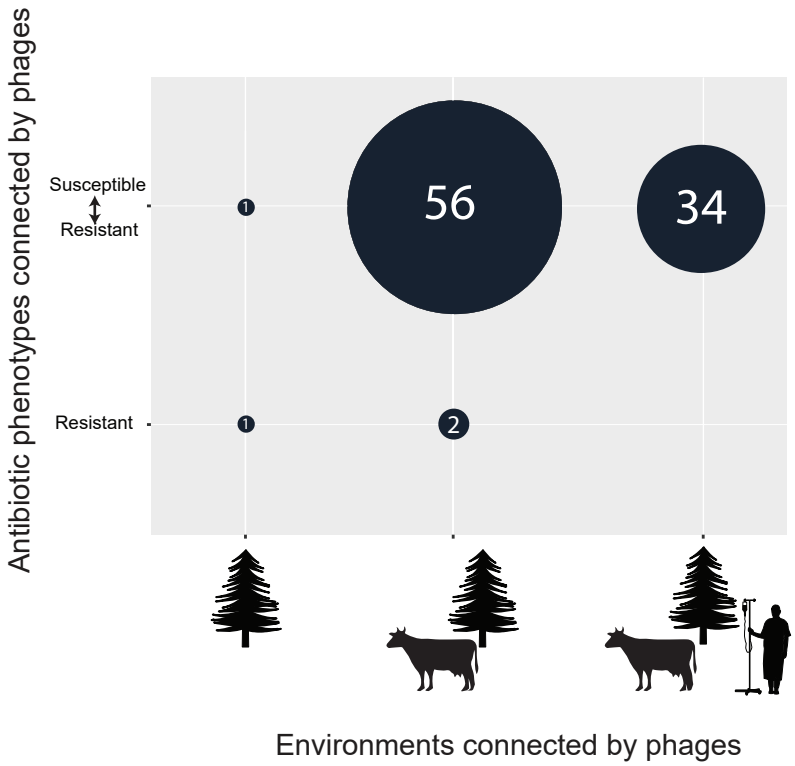
Figure 5



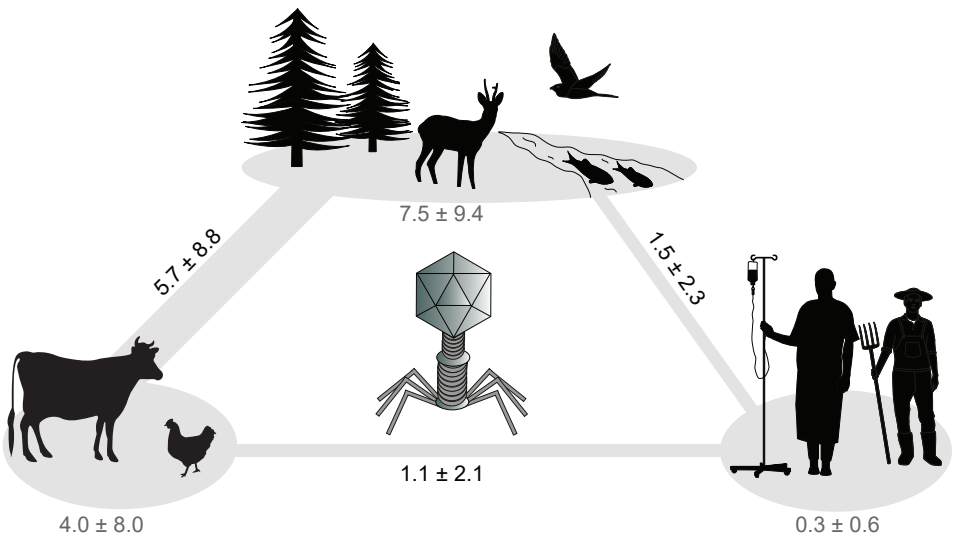
Phage susceptibility of strains classified after drug resistance phenotypes. For each strain, the number of total phages infecting this host was determined and depicted on the y-axis. On the x-axis, strains are classified according to their antimicrobial resistance phenotype. The average number of phages infecting a host phenotype is illustrated in red (mean±sd). There is no significant difference in the infection of antimicrobial susceptible or resistant hosts (Two-sided Wilcoxon rank sum test with continuity correction, W = 568.5, n = 60, p-value = 0.068).

Figure 6

(a)

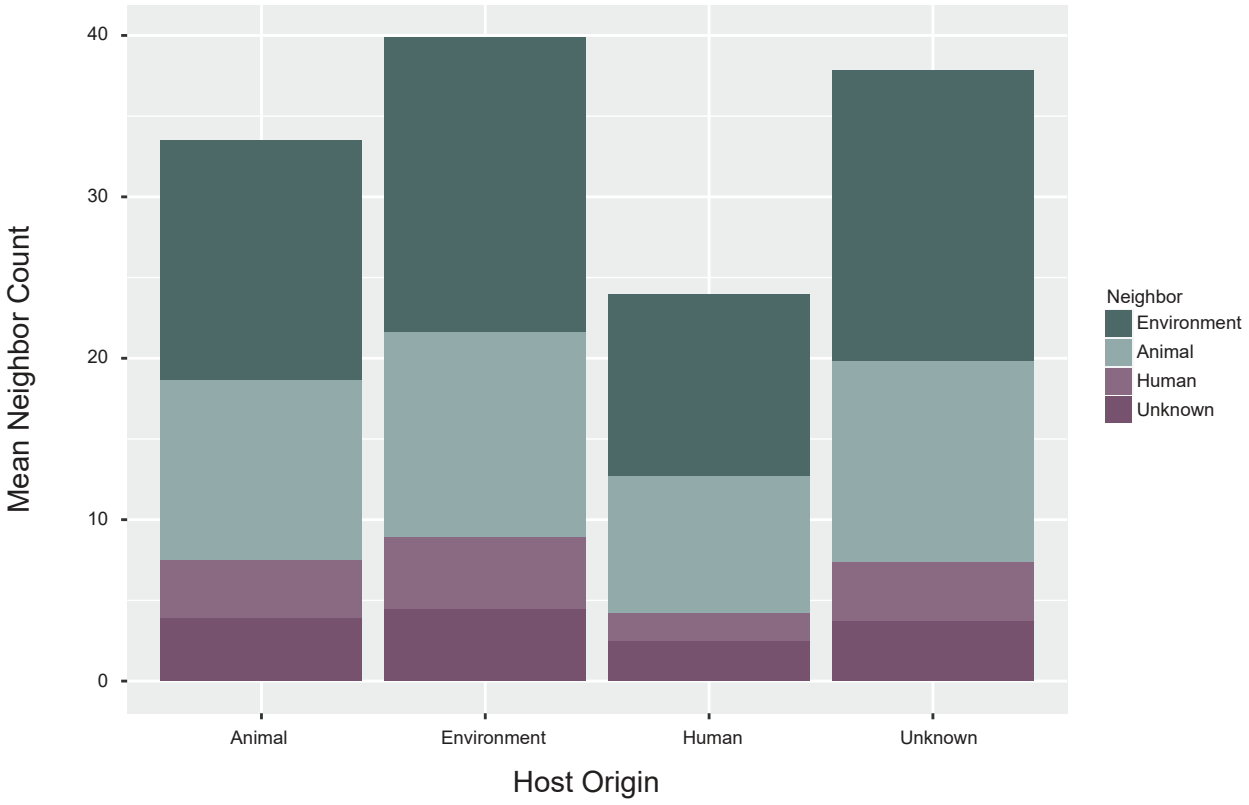


(b)



Phages connecting hosts of different antimicrobial resistant phenotypes and epidemiologic backgrounds. (a) Phages infecting either exclusively drug resistant, or drug resistant and susceptible bacteria are represented as rows. The number of phages connecting hosts from the environmental ecosystem, both the environmental and veterinary ecosystem, or all three ecosystems are represented as columns. (b) The average number of shared phages (mean±sd) between hosts within or across an ecosystem is depicted. Staphylococcal hosts are classified according to their isolation origin into environmental, veterinary, or human associated strains.

Figure 7



Mean neighbor count of a phage permissive host in the bipartite network projection. Phage permissive hosts (60) were categorized after their isolation origin, and the number of direct neighbors connected through phages was counted. Neighbors themselves were subdivided according to their isolation origin. On average, hosts isolated from animals revealed to have 33.5 ± 14.5 (n = 53) neighbors, environmental hosts 39.9 ± 11.6 (n = 23), hosts isolated from the human biome 24 ± 9 (n = 40), and hosts of unknown isolation origin 37.8 ± 12.6 (n = 7) neighbors.

Supplementary Table 2

Summary of the constitution and phage isolation efficiency for each enrichment cocktail.

Cocktail	Strains Origin	# Strains	# Isolated Phages
A	Randomly combined	9	14
B	Animal	9	32
C	WWTP/surface water	8	54
D	Labstrains	11	26
E	WWTP/surface water	9	29

Supplementary Table 4

The number of isolated phages for enrichment species, and their corresponding number of successful and unsuccessful enrichment strains.

Enrichment Species	# Enrichment Strains	# of successfull Enrichment Strains	# Enriched Phages	Enriched Phages (%)
S. aureus	14	2	10	6.45%
S. caprae/capitis	1	1	9	5.81%
S. chromogenes	2	1	4	2.58%
S. devriesei	1	0	0	0.00%
S. epidermidis	6	5	56	36.13%
S. equorum	1	1	3	1.94%
S. fleuretti	1	1	1	0.65%
S. haemolyticus	3	1	4	2.58%
S. lentus	1	1	6	3.87%
S. pseudintermedius	1	1	1	0.65%
S. saprophyticus	1	1	1	0.65%
S. schleiferi	1	1	8	5.16%
S. sciuri	4	3	21	13.55%
S. simulans	1	0	0	0.00%
S. succinus	2	2	9	5.81%
S. vitulinus	3	3	19	12.26%
S. xylosus	3	2	3	1.94%
		26	**155**	

Supplementary Table 5

Compilation of isolation hosts for all induced phages, and their respective efficiency.

Species	Strain	# Isolated Phages
S. epidermidis	C3910	2
S. epidermidis	C6869	5
S. epidermidis	I0515	6
S. epidermidis	I0564	4
S. epidermidis	NCC100655	5
S. sciuri	C6888	2

Supplementary Table 7

Summary of all phage isolation and discrimination advances on each staphylococcal species.

Species	# Successfull Isolation Strains	# Isolated phages	# Different phages
S. aureus	2	10	5
S. caprae/capitis	1	9	1
S. chromogenes	1	4	1
S. epidermidis	6	78	23
S. equorum	1	3	1
S. fleuretti	1	1	0
S. haemolyticus	1	4	3
S. lentus	1	6	3
S. pseudintermedius	1	1	1
S. saprophyticus	1	1	1
S. schleiferi	1	8	5
S. sciuri	3	23	8
S. succinus	2	9	7
S. vitulinus	3	19	14
S. xylosus	2	3	3
		179	**76**

Supplementary Table 12

Taxonomic diversity of species and strains detected in each module after modularity sorting of the phage-bacteria interaction matrix.

Module	Species Groups	Species	Number of Strains
1	Hyicus-Intermedius	*S. schleiferi*	1
1	Epidermidis-Aureus	*S. caprae/capitis*	1
1	Epidermidis-Aureus	*S. epidermidis*	7
1	Epidermidis-Aureus	*S. lugdunensis*	1
1	Epidermidis-Aureus	*S. warneri*	1
1	Saprophyticus	*S. nepalensis*	1
1	Saprophyticus	*S. pettenkoferi*	1
1	Sciuri	***S. sciuri***[1]	1
2	Epidermidis-Aureus	***S. aureus***	1
2	Epidermidis-Aureus	***S. haemolyticus***	1
2	Epidermidis-Aureus	*S. pasteuri*	1
2	Saprophyticus	*S. cohnii*	1
2	Saprophyticus	*S. kloosii*	1
2	Saprophyticus	*S. saprophyticus*	1
2	Saprophyticus	***S. xylosus***	1
2	Sciuri	*S. fleuretti*	1
2	Sciuri	***S. lentus***	2
2	Sciuri	***S. sciuri***	5
2	Sciuri	*S. vitulinus*	3
3	Epidermidis-Aureus	***S. haemolyticus***	1
3	Epidermidis-Aureus	*S. hominis*	1
3	Saprophyticus	*S. arlettae*	1
3	Sciuri	***S. lentus***	1
4	Hyicus-Intermedius	*S. chromogenes*	1
4	Hyicus-Intermedius	*S. hyicus*	1
4	Hyicus-Intermedius	*S. pseudintermedius*	3
4	Epidermidis-Aureus	***S. aureus***	6
4	Epidermidis-Aureus	*S. capitis*	1
4	Auricularis	*S. auricularis*	1
4	Saprophyticus	*S. equorum*	1
4	Saprophyticus	*S. succinus*	2
4	Saprophyticus	***S. xylosus***	8

[1] species in bold occur in one than 1 module

Figures

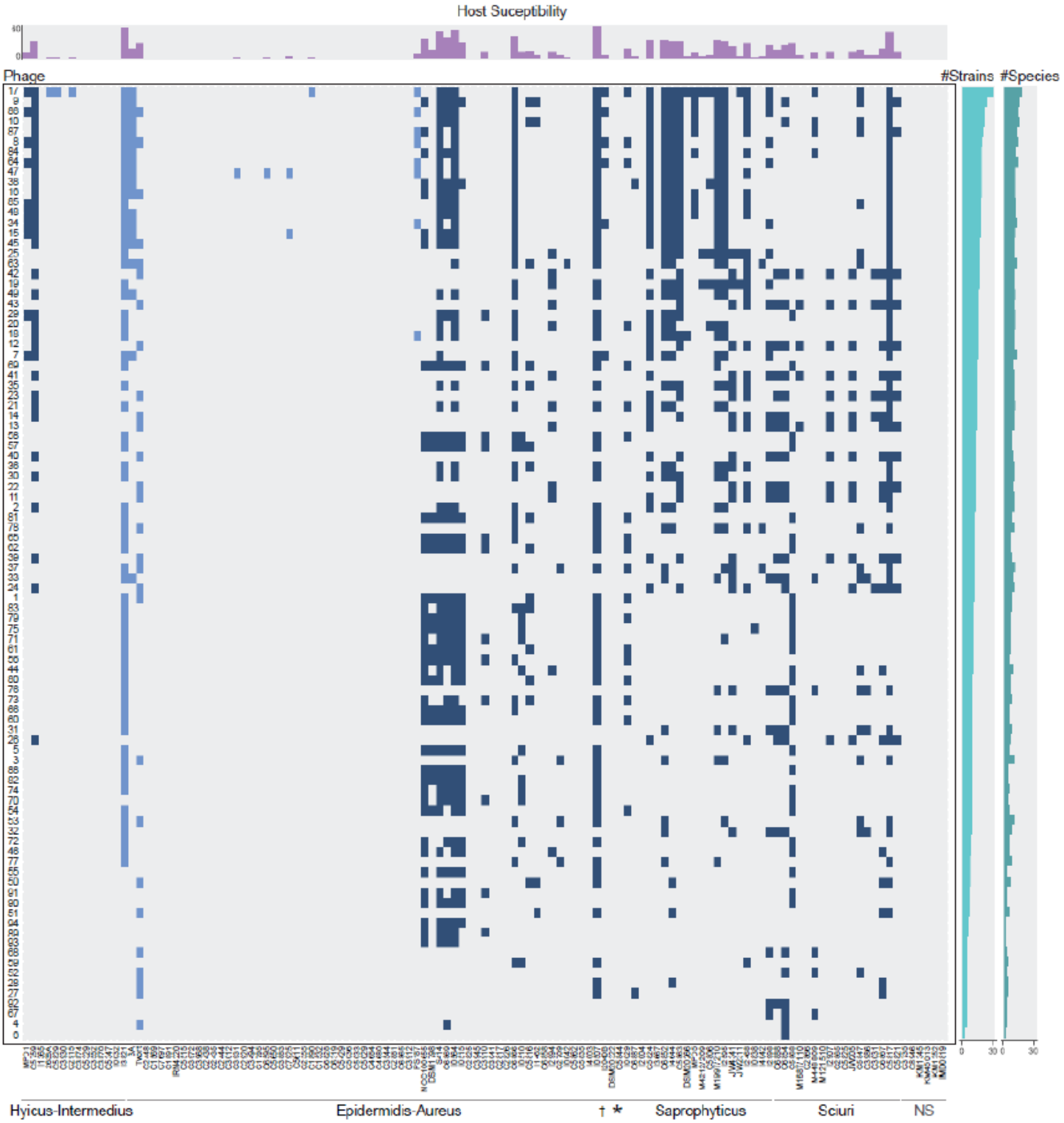


Figure 1

A staphylococcal phage-bacteria incidence matrix. Bacterial lawns of 123 hosts from 32 species, were challenged with 94 different staphylococcal phages from wastewater and phage K. Phages on the y-axis are sorted from broad host range to narrow. Bacterial hosts in columns are sorted after cluster-groups

and subdivided species as established in 17,18. †: Species-group Auricularis. *: cluster-group Simulans. NS: Non-Staphylococcus hosts. Each blue-colored square of the incidence matrix corresponds to a phage-host infection where single plaques were visible. Squares in dark blue indicate infections on CoNS and squares in light blue on CoPS. The phage permissiveness for each host is indicated in the host susceptibility bar chart on top of the incidence matrix, which represents the number of phages infecting a strain. The two bar charts on the right indicate the total number of strains (# strains, left) and species (# species, right) a phage infected. The incidence matrix has a diameter of six and a density of 0.1 (=1135/11562). Phages are abbreviated with their final unique numerical identifier (PG- 2021_*).

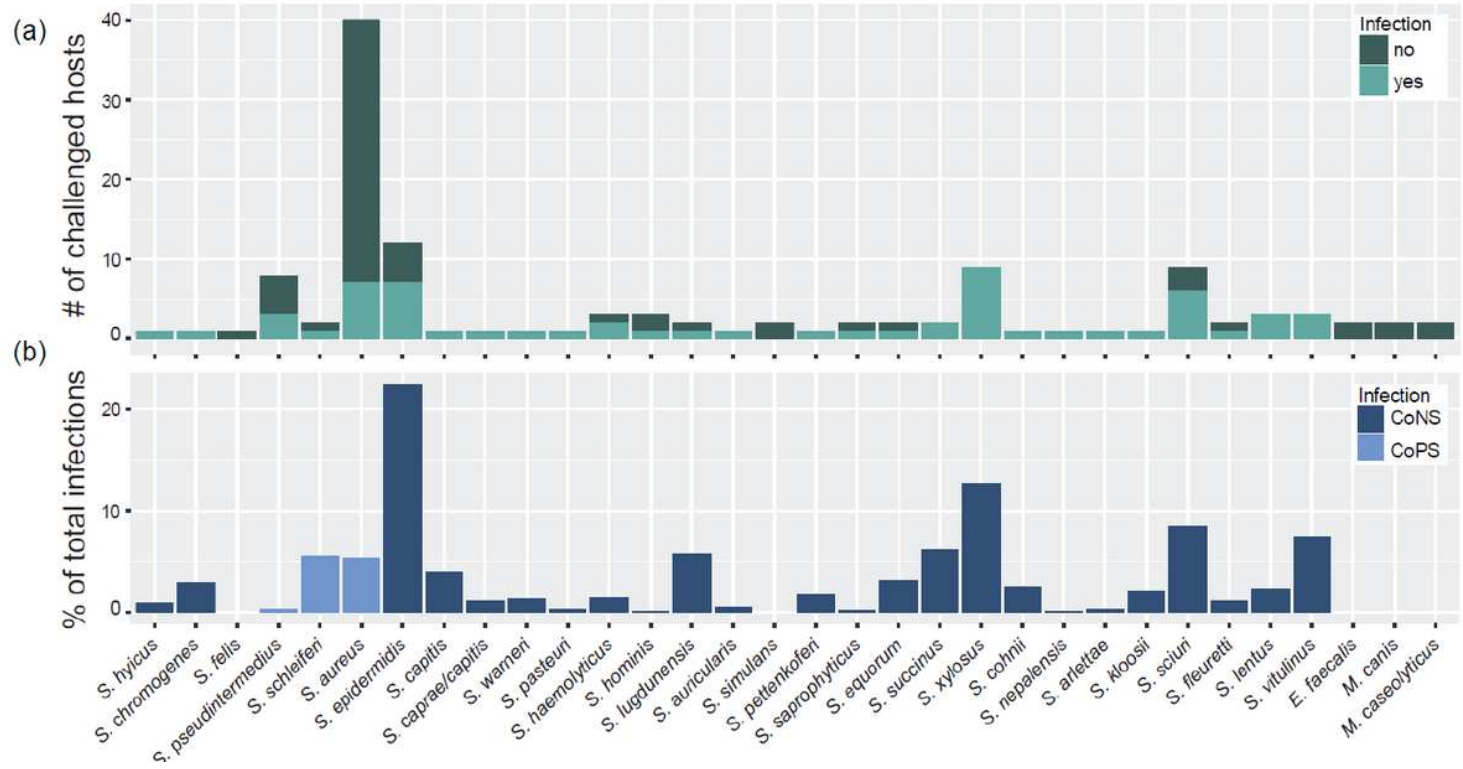


Figure 2

Phage infections on staphylococcal species. Species challenged in the phage-bacteria interaction matrix are shown on the x-axis and sorted after the established Staphylococcus species groups18. (a) For each species, the number of phage resistant and susceptible strains are depicted. (b) Phage infections on each respective species was plotted as a percentage of the total infections detected in the interaction matrix.

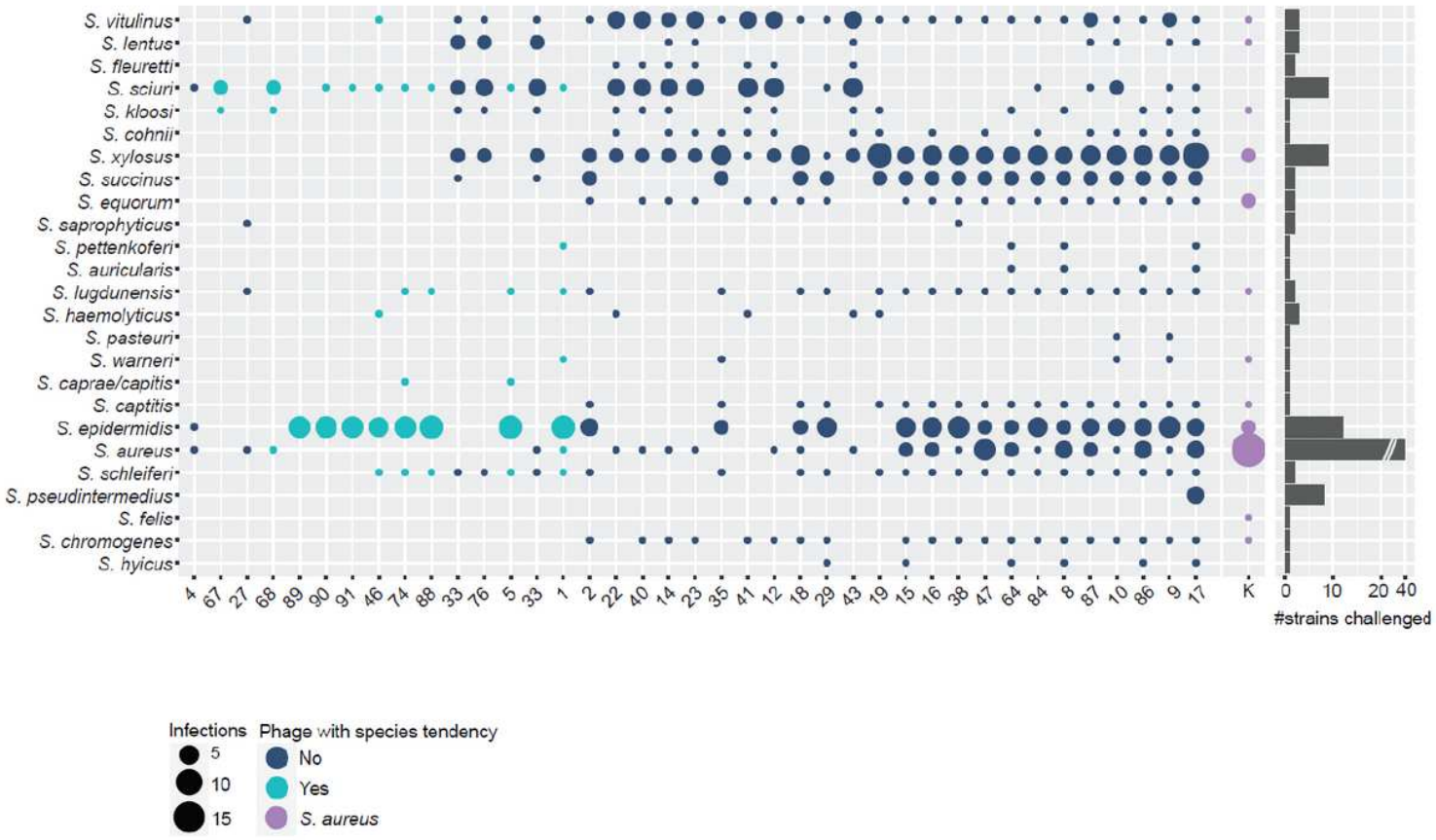


Figure 3

Illustration of the host ranges collapsed on the species level for the 40 sequenced phages. Phages on the x-axis are sorted from narrow (left) to broad host range (right). Species on the y-axis are sorted after phylogenetic relationship in species groups18. A phage host range is depicted as a column, where infection of a staphylococcal species is illustrated using circles. For each respective species, the area of the circle is scaled according to the number of strains a phage can replicate on (scale: 1-15). The total number of strains challenged per species is depicted in the bar-chart on the right. Host ranges on this host array are colored as follows: Phage with species tendency (≥ 50 % of all infections on a single species) in turquoise; phages with no clear species tendency in dark blue; polyvalent phage K in violet. Phages are abbreviated with their final unique numerical identifier (PG-2021_*).

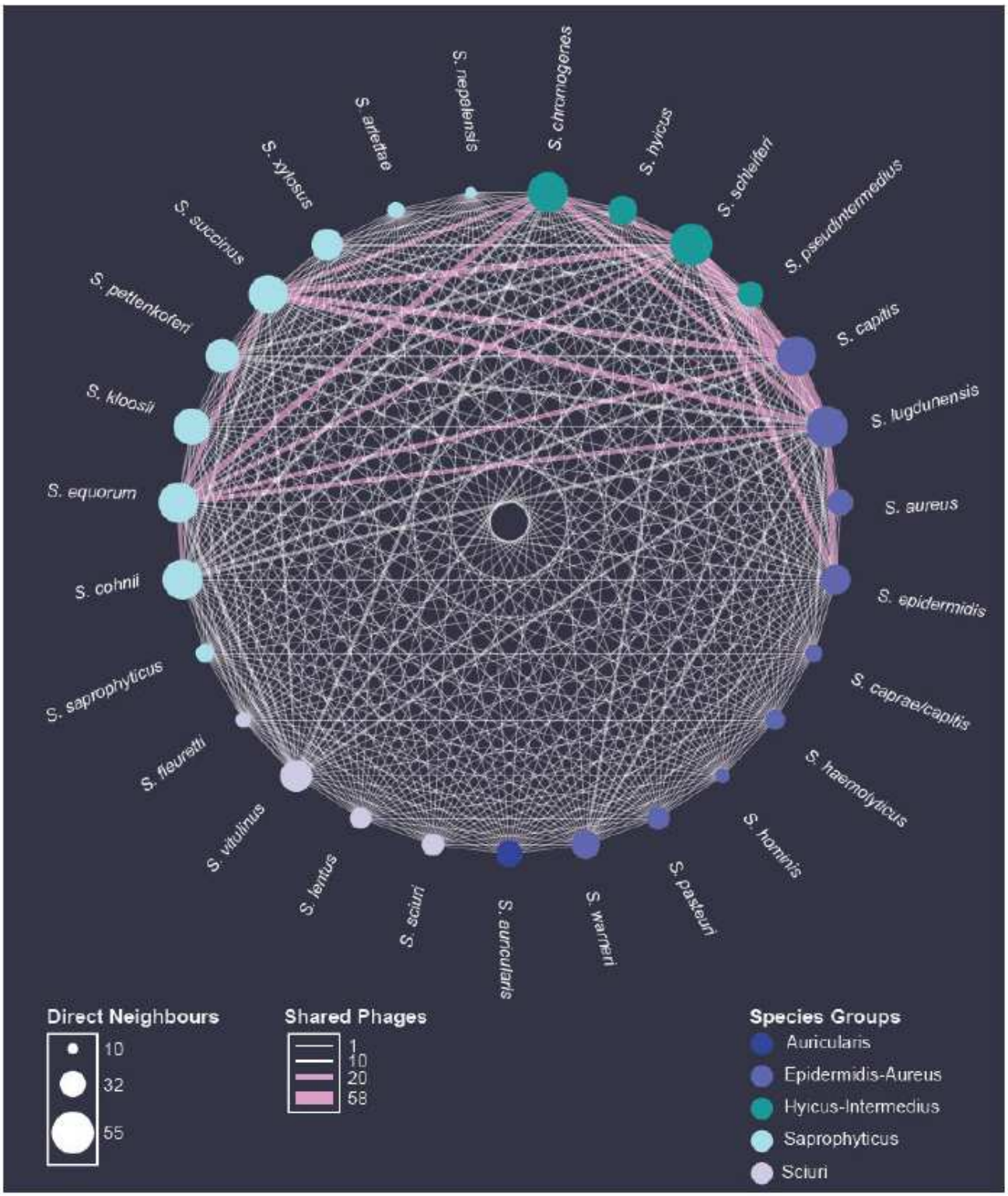


Figure 4

Species network with phages as coupling links. Staphylococcal host species are represented as nodes and sorted after cluster affiliation18. The area of each node directly correlates with the average number of strain neighbors a species is connected to. The number of shared phages between species is represented as weighted edges. If >20 phages are shared between two staphylococcal species, edges are colored in pink.

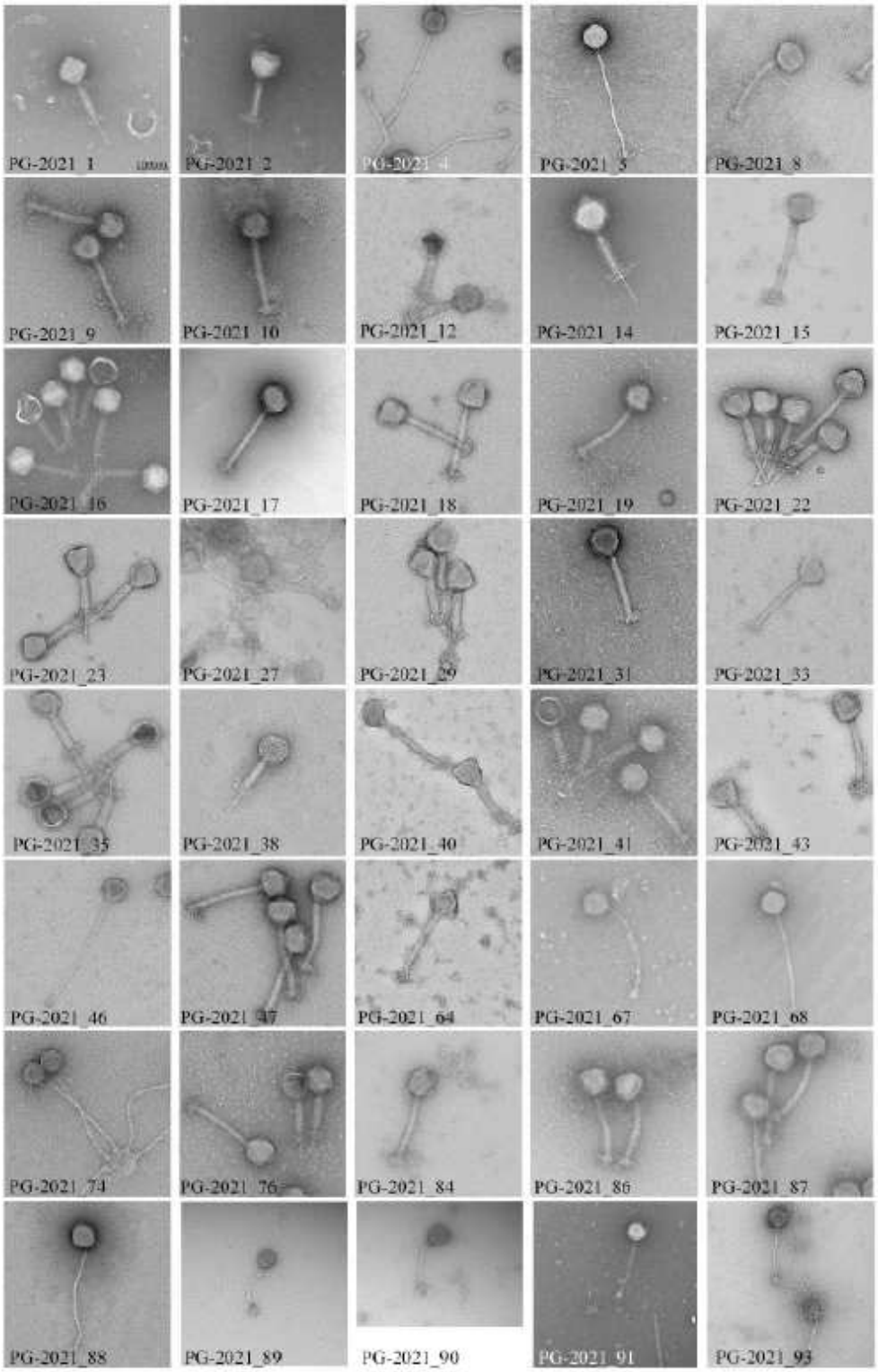


Figure 5

Electron micrographs of the sequenced staphylococcal phages. Phages were isolated from the wastewater treatment plant inlet, outlet, or by induction of bacterial lysogens. All pictures are adjusted according to the displayed scale-bar on the top-left corner.

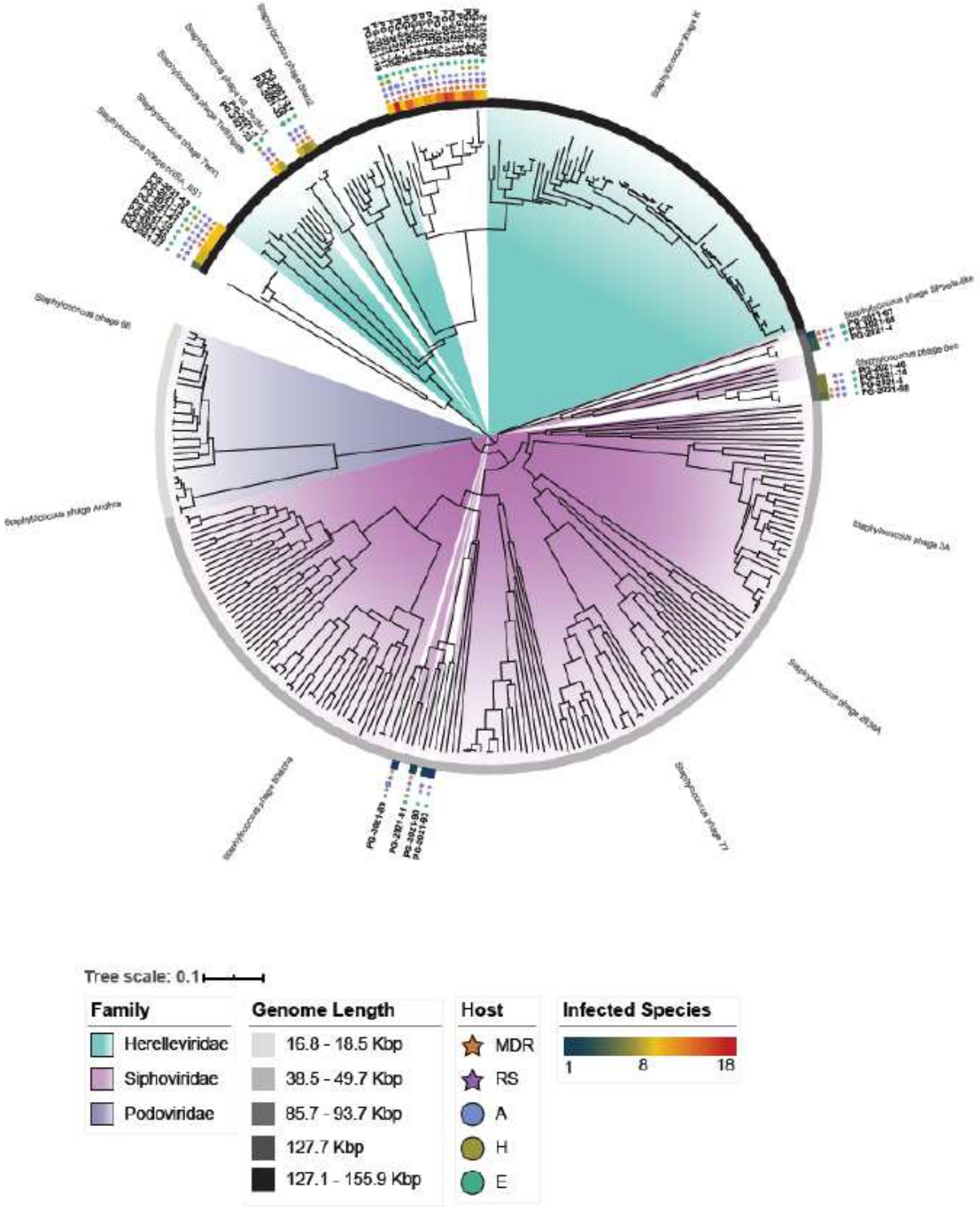


Figure 6

Phylogenomic tree of staphylococcal phages. All published staphylococcal phages (Supplementary Table 15) are displayed together with the here isolated and sequenced CoNS infecting viruses. For each phage genus, a representative phage is indicated. Phages from our collection are represented in bold and their corresponding host range is represented as follows: the number of infected species is indicated using a continued color scale; isolation origin, and antimicrobial resistant phenotype of infected hosts are

represented using colored circles and stars, respectively. Phages infect A: hosts isolated from animals, H: hosts isolated from humans, E: hosts isolated from the environment. MDR: host is multidrug resistant. RS: phage infects hosts with antimicrobial resistant and susceptible phenotypes.

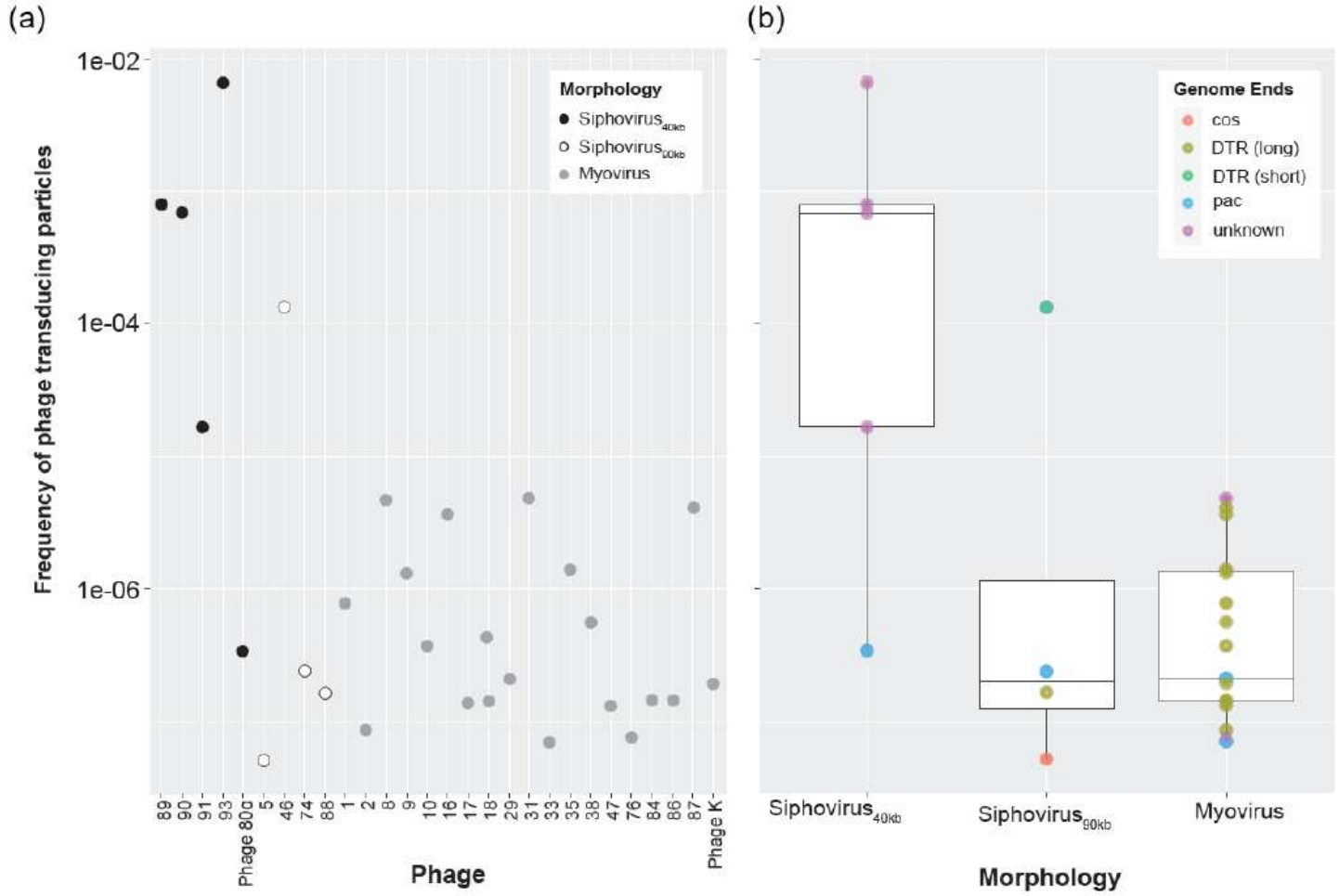


Figure 7

Frequency of transducing particles for diverse staphylococcal phages. (a) Estimated frequency of transducing particles for each respective phage and corresponding phage morphology. Phages are abbreviated with their final unique numerical identifier (PG-2021_*). (b) Mean frequencies of transducing particles for each phage morphology. Phage termini were detected using PhageTerm96 and are illustrated using colors. DTR: direct terminal repeats.