

# SCIENTIFIC REPORTS



OPEN

## Safety assessment of *Staphylococcus* phages of the family *Myoviridae* based on complete genome sequences

Received: 28 June 2016  
Accepted: 19 December 2016  
Published: 24 January 2017

Zelin Cui<sup>1</sup>, Xiaokui Guo<sup>2</sup>, Ke Dong<sup>2</sup>, Yan Zhang<sup>2</sup>, Qingtian Li<sup>3</sup>, Yongzhang Zhu<sup>2</sup>, Lingbing Zeng<sup>2</sup>, Rong Tang<sup>1</sup> & Li Li<sup>1</sup>

*Staphylococcus* phages of the *Myoviridae* family have a wide host range and potential applications in phage therapy. In this report, safety assessments of these phages were conducted based on their complete genome sequences. The complete genomes of *Staphylococcus* phages of the *Myoviridae* family were analyzed, and the Open Reading Frame (ORFs) were compared with a pool of virulence and antibiotic resistance genes using the BLAST algorithm. In addition, the lifestyle of the phages (virulent or temperate) was also confirmed using PHACTS. The results showed that all phages were lytic and did not contain resistance or virulence genes based on bioinformatic analyses, excluding the possibility that they could be vectors for the dissemination of these undesirable genes. These findings suggest that the phages are safe at the genome level. The SceD-like transglycosylase, which is a biomarker for vancomycin-intermediate strains, was widely distributed in the phage genomes. Approximately 70% of the ORFs encoded in the phage genomes have unknown functions; therefore, their roles in the antibiotic resistance and virulence of *Staphylococcus aureus* are still unknown and require consideration before use in phage therapy.

Due to the appearance and dissemination of antibiotic-resistant bacteria, concern is increasing that certain bacteria will develop resistance to all known antibiotics, and we will enter a post-antibiotic era. Recently, phage therapy has regained interest among scientists, and some phages have been used in clinical trials with no harmful effects on the trial participants<sup>1–3</sup>.

While phages can be vectors for horizontal gene transfer, several studies have shown that they can also be sources for the transmission of virulence and antibiotic resistance genes among bacteria<sup>4</sup>, thereby accelerating the evolution of virulence and/or antibiotic resistance in bacteria. Phages containing such undesirable genes could be harmful to human health and must be excluded from phage therapy. One study confirmed that Stx production was phage-regulated in stx(2g)-positive strains (Shiga toxin-producing *Escherichia coli* (STEC))<sup>5</sup>. Report also showed that bacteriophages from poultry, cattle and pigs containing antibiotic resistance genes<sup>6</sup>. These studies suggest that bacteriophages could be environmental vectors for the horizontal transfer of virulence and antibiotic resistance genes.

*Staphylococcus* phages of the *Myoviridae* family have been isolated worldwide. Published data show that they have a wide host range and can potentially be used for phage therapy. To date, the complete genome sequences of dozens of *Staphylococcus* phages of the *Myoviridae* family have become available, including phage K, G1, Twort, A5W, Sb-1, ISP, SA5, GH15, JD007, SA11, vB\_SauM\_Remus, vB\_SauM\_Romulus, S25-3, S25-4, phiPLA-RODI, phiPLA-C1C, phiSA012, Team1, P108, MCE-2014, 812, SA1, Staphy1N, MSA6, 676Z, P4W, and Fi200w<sup>7–15</sup>. These phages can potentially be used for the treatment of infectious diseases caused by *Staphylococcus aureus* in animal models. Phage S25-3 has been shown to prolong life in the silkworm larval infection model and other animal infection models<sup>16</sup>. The *Staphylococcus* phages Romulus and Remus have infected approximately 70% of the

<sup>1</sup>Department of Laboratory Medicine, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, 200080, China. <sup>2</sup>Department of Immunology and Microbiology, School of Medicine, Shanghai Jiao Tong University, Shanghai, 200025, China. <sup>3</sup>Department of Laboratory Medicine, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, China. Correspondence and requests for materials should be addressed to Z.C. (email: czl@sjtu.edu.cn) or L.L. (email: annylish@126.com)

tested *S. aureus* strains and display promising lytic activity against these isolates<sup>15</sup>. The *Staphylococcus* phage ISP was shown to be effective against 86% of tested isolates, including relevant methicillin-resistant *S. aureus* (MRSA) strains<sup>7</sup>. These studies demonstrate the potential use of phages for the prevention of infectious diseases caused by *S. aureus*. In the present report, a safety assessment based on complete genome sequences of *Staphylococcus* phages of the *Myoviridae* family was conducted to evaluate their suitability for clinical use.

## Methods

**Complete phage genomes.** *Staphylococcus* phages of the *Myoviridae* family, including the previously reported phage JD007, had complete genome sequence data. The complete genome sequences were collected from the NCBI database (<http://www.ncbi.nlm.nih.gov/genome/viruses/>). All *Staphylococcus* phages belonged to the *Myoviridae* family and were verified by detailed information provided by the NCBI phylogeny system and their morphology. The published papers related to these phages were also screened. The *Staphylococcus* phages were confirmed to belong to the *Myoviridae* family by morphological characterization of the contractile sheath between the head and tail.

**General genomic features of the phages.** The complete genome sequences of *Staphylococcus* phages of the *Myoviridae* family were downloaded in FASTA format from NCBI along with their annotation information. The genomes of the phages were re-annotated using RAST, and each ORF and protein domain search was performed using the InterProScan program and CDD<sup>17,18</sup>. The phylogeny tree was constructed using MEGA5 with the complete genome sequences based on the neighbour-joining method<sup>19</sup>. The visual alignments of the complete genomes were constructed using the Mauve2.3.1 program<sup>20</sup>.

**Lifestyles of the phages.** The lifestyles of phages (virulent or temperate) were predicted using the PHACTS program (<http://www.phantome.org/PHACTS/index.htm>)<sup>21</sup>. The genome annotation results of each phage were analyzed using PHACTS, and their lifestyle was confirmed. PHACTS utilizes a novel algorithm and a supervised Random Forest classifier to predict whether the lifestyle of a phage is virulent or temperate. The algorithm creates a training set from phages with known lifestyles. PHACTS predictions have been shown to have a 99% precision rate, and PHACTS can also determine the lifestyle of a phage using only genomic data. A current limitation of PHACTS is that a confident lifestyle prediction cannot be made for a small number of phages. This is primarily due to the variability that arises from the random sampling during classifications. If an unknown phage does not have any similarity to phages with known lifestyles in the database, the predictions will be less certain. It is expected that as more phages with known lifestyles are added to the database, the precision rate and sensitivity of the predictions will increase<sup>21</sup>. Additionally, the experimental results of the bacteriophage one step growth curve reported in the corresponding papers were also analyzed to confirm the phage lifestyles.

**Analysis of virulence and antibiotic resistance genes.** All annotated phage ORFs were used to search an antibiotic resistance gene database (ARDB, <http://ardb.cbcb.umd.edu/>)<sup>22</sup> and a virulence factor database (VFDB, <http://www.mgc.ac.cn/VFs/main.htm>)<sup>23</sup>. Hits with more than 70% coverage and 30% identity were considered as positive results. Lihong *et al.* improved the infrastructural dataset of VFDB: (i) removed the redundancy introduced by previous releases and generated two hierarchical datasets—one core dataset of experimentally verified virulence factors (VFs) only and another full dataset including all known and predicted VFs. Their efforts enhanced the data quality of the VFDB and promoted the usability of the database in the big data era for the bioinformatic mining of the explosively growing data regarding bacterial VFs<sup>23</sup>. Antibiotic Resistance Genes Database (ARDB)—unifying most of the publicly available information on antibiotic resistance. Each gene and resistance type is annotated with rich information, including resistance profile, mechanism of action, ontology, COG and CDD annotations, as well as external links to sequence and protein databases. ARDB database also supports sequence similarity searches and implements an initial version of a tool for characterizing common mutations that confer antibiotic resistance. The information they provided can be used as compendium of antibiotic resistance factors as well as to identify the resistance genes of newly sequenced genes, genomes, or metagenomes<sup>22</sup>.

**Phylogenetic relationship of the phages.** To identify the phylogenetic relationship among *Staphylococcus* phages of the *Myoviridae* family, other *S. aureus* phages, including *Podoviridae* (including *Staphylococcus* phages 44AHJD, 66, GRCS, P68, PT1028, S24-1, and SAP-2) and *Siphoviridae* (including *Staphylococcus* phages 3MRA, 13, 29, 52A, 55, 69, 77, 92, 187, 2638A, DW2, IME-SA4, Ipla7, P954, phi5967PVL, phiSa119, phiSauS-IPLA88, PVL, SA13, SA97, SAP-26, SpaA1, StauST-398-3, StB20, StB20-like, StB27, 3A, 11, 23MRA, 37, 71, 88, 96, B166, B236, CNPH82, Ipla5, JS01, phiBU01, phiETA, phiETA2, phiETA3, phiJB, phiMR11, phiNM1, and phiNM2) (phages' nucleic acid access numbers refer additional file 1), were chosen to evaluate the phylogenetic relationship using complete genome sequences (<http://www.ncbi.nlm.nih.gov/genome/viruses/>). The neighbour-joining tree was constructed using MEGA5<sup>19</sup>.

## Results

**General genomic features of *Staphylococcus* phages of the *Myoviridae* family.** A total of 22 *Staphylococcus* phages with complete genome sequence data were examined: K, G1, Twort, A5W, Sb-1, ISP, SA5, GH15, JD007, SA11, vB\_SauM\_Remus, vB\_SauM\_Romulus, S25-3, S25-4, phiIPLA-RODI, phiIPLA-C1C, phiSA012, Team1, P108, MCE-2014, 812, and SA1<sup>7-15</sup>. The *Staphylococcus* phages Staphy1N, MSA6, 676Z, P4W, and Fi200w, for which only partial genome sequences were available, were excluded from this study. The phages have been isolated worldwide, and all belong to the *Myoviridae* family. As shown in Table 1, the genomes were approximately 127 kb–148 kb in length, contained between 170 and 220 ORFs, and had a GC content of 27–31%. However, the GC content of phage SA1 was 45.83%, which was much higher than other *Staphylococcus* phages of the *Myoviridae* family.

Phages	Country	Size (bp)	GC %	Access No.	Lytic/Lysogenic	Date	Refs
K	USA	127395	30.60	NC_005880	lytic	2004	10
G1	Georgia	138715	30.39	AY954969.1	lytic	2005	33
Twort	Canada	130706	30.26	NC_007021	lytic	2005	33
A5W	Poland	137083	30.47	EU418428.1	lytic	2008	NCBI
Sb-1	Georgia	127188	30.48	HQ163896.1	lytic	2011	34
ISP	Georgia	138339	30.42	FR852584.1	lytic	2011	7
SA5	Georgia	137031	30.42	JX875065	lytic	2012	NCBI
GH15	China	139806	30.23	JQ686190	lytic	2012	14
JD007	China	141836	30.37	JX878671	lytic	2012	13
SA11	Korea	136326	30.04	JX194239	lytic	2012	8
vB_SauM_Remus	Belgium	134643	29.97	JX846612	lytic	2013	15
vB_SauM_Romulus	Belgium	131332	30.01	JX846613	lytic	2013	15
S25-3	Japan	139738	30.22	AB853330	lytic	2013	9
S25-4	Japan	132123	30.31	AB853331	lytic	2013	9
phiPLA-RODI	Belgium	142348	30.42	KP027446	lytic	2015	11
phiPLA-C1C	Belgium	140961	27.98	KP027447	lytic	2015	11
phiSA012	Japan	142094	30.31	NC_023573	lytic	2009	35
Team1	Canada	140903	30.33	KC012913.1	lytic	2014	36
P108	China	140807	30.22	KM216423.1	lytic	2014	NCBI
MCE-2014	U.K.	141907	30.38	NC_025416	lytic	2014	37
812	USA	142096	30.40	NC_029080	lytic	2016	NCBI
SA1	USA	147303	45.83	NC_027991	lytic	2015	NCBI

**Table 1.** The *Staphylococcus* phages of the *Myoviridae* family with published complete genome sequences (through June, 2016).

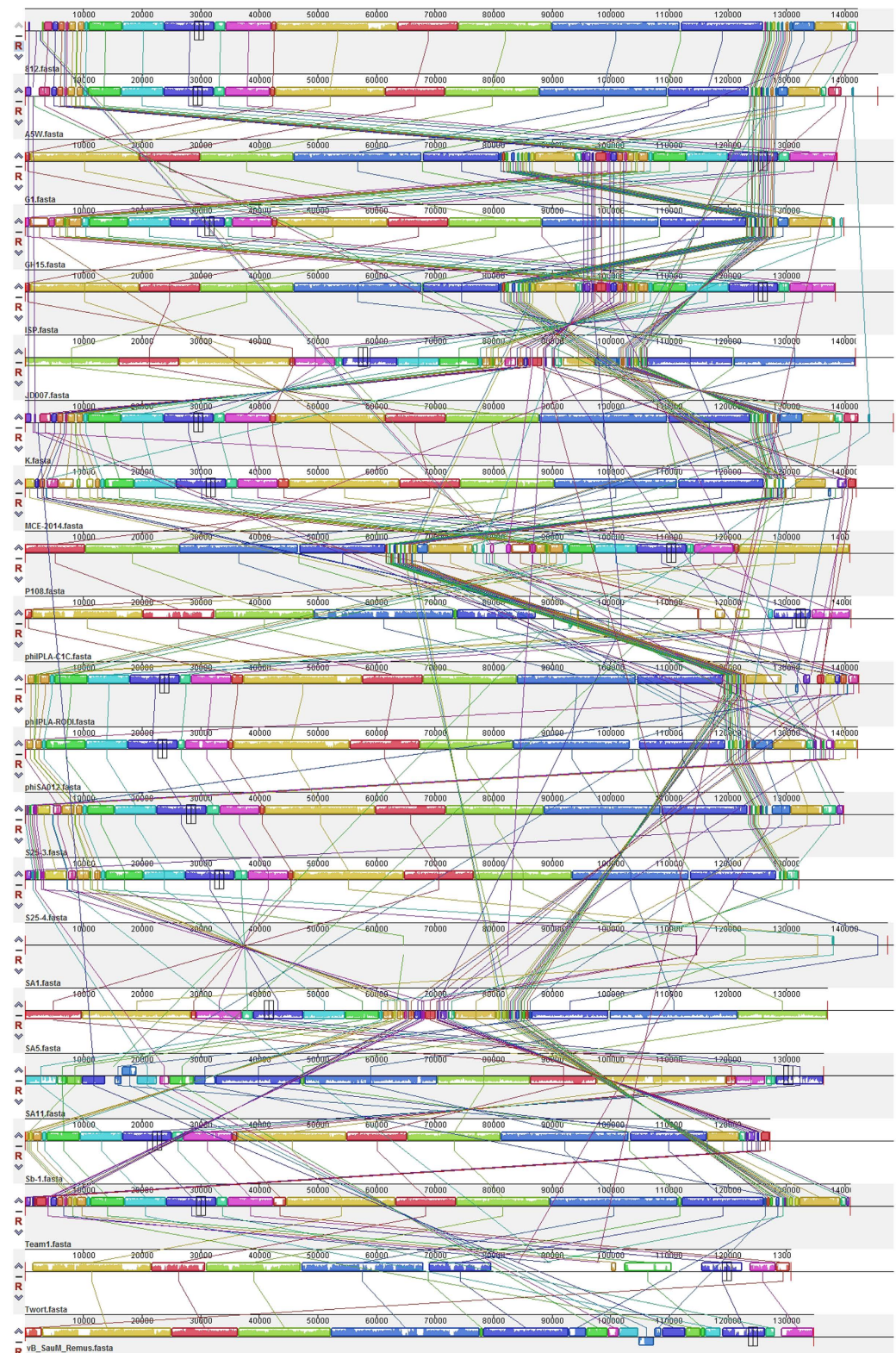
**Lytic or lysogenic analysis.** PHACTS, a computational approach used to classify the lifestyle of bacteriophages, was used after the complete genome sequences were annotated and the proteins were converted into the FASTA format. The results showed that the *Staphylococcus* phages K, G1, Twort, A5W, Sb-1, ISP, SA5, GH15, JD007, SA11, vB\_SauM\_Remus, vB\_SauM\_Romulus, S25-3, S25-4, phiPLA-RODI, phiPLA-C1C, phiSA012, Team1, P108, MCE-2014, 812, and SA1 belong to the *Myoviridae* family and are all lytic phages<sup>7–15</sup>. The previous report shows that phage phiPLA-RODI and phiPLA-C1C are lytic phages<sup>11</sup>, our results were consistent with the studies reported previously.

**Genome organization and annotation.** The complete genomic sequences of *Staphylococcus* phages of the *Myoviridae* family were compared. As showed in Fig. 1, the genomic structure was chimeric and linear. Several functional modules shown in Table 2 consistently exhibited mosaicism in the genomes, including those involved in packaging, lysis, structure, phage and host interactions, DNA manipulation, and some ORFs with additional functions. High degrees of sequence similarity were evident across the phage sequences, which is unique among *Staphylococcus* phages of the *Myoviridae* family. While, *recombinase* gene was widely existed in the complete genomes of *Staphylococcus* phages of the *Myoviridae* family, what suggest that genomic sequences combination may occur between the homologue sequences during phage infection. The difficulties to knock out *recombinase* gene in these phages may block experimental validation the fully inactivity of Recombinase or confirmation there were no homologue sequences combination occur in the *recombinase* knock-out phages.

**Analysis of virulence and antibiotic resistance genes.** All of the predicted ORFs were compared with the sequences in the Antibiotic Resistance Genes Database (ARDB) and the virulence factors database VFDB using the BLASTn algorithm, with the criteria that genes with more than 70% coverage and 30% identity were considered to exhibit positive results. No significant hits based on the ORFs predicted in the complete genomes of the *Staphylococcus* phages of the *Myoviridae* family were obtained.

As shown in Fig. 2, the SceD-like transglycosylase, which is a biomarker for vancomycin-intermediate strains<sup>24</sup>, was encoded in the genome of *Staphylococcus* phages K, G1, Twort, A5W, Sb-1, ISP, SA5, GH15, JD007, SA11, vB\_SauM\_Remus, vB\_SauM\_Romulus, S25-3, S25-4, phiPLA-RODI, phiPLA-C1C, phiSA012, Team1, P108, MCE-2014, and 812<sup>7–15</sup>. Furthermore, approximately 70% of the ORFs were predicted to have unknown functions. It suggested that these ORFs encoding proteins with no similarity functional domains of proteins in InterProScan and CDD database.

**Clustering of *Staphylococcus* phages of the *Myoviridae* family.** The phylogeny tree showed that *Staphylococcus* phages of the *Myoviridae* family clustered in the same branch, based on the alignments of the complete genome sequences. As shown in Fig. 3, the phages K, G1, Twort, A5W, Sb-1, ISP, SA5, GH15, JD007, SA11, vB\_SauM\_Remus, vB\_SauM\_Romulus, S25-3, S25-4, phiPLA-RODI, phiPLA-C1C, phiSA012, Team1, P108, MCE-2014, 812, and SA1<sup>7–15</sup> clustered together on the same sub-branch with other *Staphylococcus* phages of the *Myoviridae* family. These results are consistent with the morphology of the phages. It is well known that

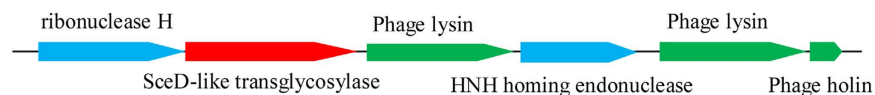


**Figure 1.** Comparative genomic analysis of *Staphylococcus* phages of the *Myoviridae* family.

the proteins such as Helicase, DNA Polymerase, Primase, Large terminase and Major Capsid Protein encoded in the genome of phages are usually chosen to draw phylogenetic trees<sup>25</sup>. However, such genes are not always simultaneously existed in the genomes of phages our study enrolled, so it is difficult to cluster phages using these genes when some phages have while others are not. With the development of genome sequencing recently, the complete genome sequences of viruses with high identity were chosen to draw the phylogenetic trees with high accuracy.

Modules	Functional ORFs	812	A5W	G1	GH15	ISP	JD007	K	MCE-2014	P108	phiPLA-C1C	phiIPLA-ROD1	phiSA012	S25-3	S25-4	SA11	Twort	Sb-1	Team1	SA5	vB_SauM_Romulus	vB_SauM_Remus	SA1
Replication	DNA helicase	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	DNA polymerase I	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	DNA primase	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Ribonuclease H	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Introns	HNH homing endonuclease	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
Lysis	Endolysin	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Holin	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Packaging	Terminase, large subunit	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Structure	Baseplate	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Capsid and scaffold	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Major capsid protein	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Major tail protein	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Major tail sheath	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Regulation	DNA transfer protein	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	DNA-binding protein	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Integration host factor	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Recombination	Recombinase				+		+		+		+		+		+		+		+		+	+	
	Exonuclease	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Endonuclease	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	RecA protein	+	+	+		+		+		+		+		+		+		+		+		+	+

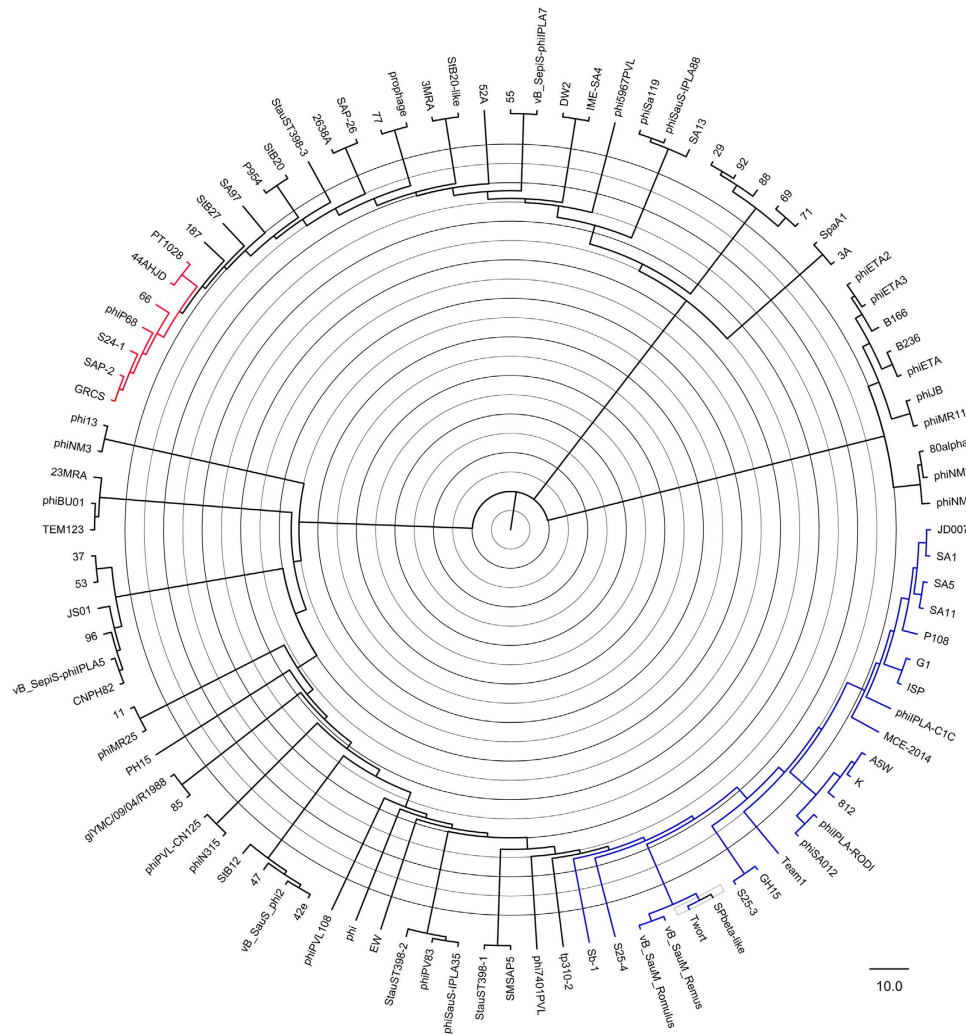
**Table 2.** The functional modules of the ORFs in the genomes of *Staphylococcus* phages of the *Myoviridae* family.



**Figure 2.** The *SceD*-like transglycosylase gene is widely distributed in the genome of *Staphylococcus* phages of the *Myoviridae* family.

## Discussions

Phage therapy has become a subject of renewed interest recently, and some phages have been used in clinical trials, including the *Staphylococcus* phages of the *Myoviridae* family. Several studies have evaluated the safety of certain phages in clinical trials. Rhoads *et al.* reported a study in which ulcers were treated for 12 weeks with bacteriophages targeting *Pseudomonas aeruginosa*, *S. aureus*, and *Escherichia coli* or a saline control<sup>3</sup>. Follow-up continued until week 24, and no adverse events were attributed to the phage therapy. In addition, no significant difference was determined between the test and control groups regarding the frequency of adverse events, rate of healing, or frequency of healing<sup>3</sup>. Unlike normal antibiotics, a bacteriophage is a type of virus that contains genetic material. Therefore, potentially useful phages require full evaluation prior to use in phage therapy. A bacteriophage may be a potential mechanism for the transmission of antibiotic resistance genes or virulence genes among bacteria. A prophage may contribute the pathogenic traits of *Enterococcus faecalis*<sup>26</sup>. Antibiotic resistance



**Figure 3.** The phylogenetic tree of *Staphylococcus* phages of the *Myoviridae* family. The diagram represents the phylogeny constructed using the complete phage genome sequences via the neighbour-joining method.

genes, including *blaTEM*, *qnrA*, *blaCTX-M-1*, and toxin genes, have been found in the DNA of bacteriophages isolated from human faecal samples<sup>27,28</sup>. Therefore, safety assessment of phages at the genome level is critically important. First, an analysis should determine whether the bacteriophage carries genes that may accelerate the virulence and antibiotic resistance of bacteria when these genes integrate into their genomes. Second, the ability of the bacteriophage genome that will integrate into the genome of bacteria should be assessed. In this evaluation, it is important and necessary to assess the safety of bacteriophages at the genome level. McCallin *et al.* evaluated the safety of a phage cocktail from one company, and a small volunteer trial did not report an association between adverse effects and oral phage exposure<sup>29</sup>. During that study, the genome sequences of phages were obtained, and bioinformatics analysis revealed that no undesirable genes were present<sup>29</sup>. A safety assessment of the phage at the genome level is an essential step during the evaluation of the suitability of phages for therapeutic applications.

*Staphylococcus* phages of the *Myoviridae* family are potentially useful for phage therapy. Dozens of phages have been shown to protect mice from death after infection with *S. aureus*. Some phages have been used in clinical trials, and no adverse effects have been reported. In this report, our results showed that none of these bacteriophages carried any antibiotic resistance or virulence genes. In addition, all of the studied bacteriophages were virulent and were therefore unable to integrate into the genome of *S. aureus*. Our studies show that *Staphylococcus* phages of the *Myoviridae* family are safe at the genome level. While, as showed in Table 2, *recombinase* was widely distributed in phage GH15, JD007, MCE-2014, phiIPLA-C1C, phiSA012, SA11, Twort, Team1, vB\_SauM\_Romulus and vB\_SauM\_Remus, it was reported that Recombinase could recombine homologous genome sequences between bacteriophages or/and its host<sup>30,31</sup>; and the *SceD*-like transglycosylase, which is a biomarker for vancomycin-intermediate strains<sup>24</sup> is encoded in the genome of *Staphylococcus* phages K, G1, Twort, A5W, Sb-1, ISP, SA5, GH15, JD007, SA11, vB\_SauM\_Remus, vB\_SauM\_Romulus, S25-3, S25-4, phiIPLA-RODI, phiIPLA-C1C, phiSA012, Team1, P108, MCE-2014, and 812<sup>7-15</sup>. The expression of this gene during *S. aureus* infection may increase the minimum inhibitory concentration of vancomycin used for the treatment of infections caused by MRSA; furthermore, approximately 70% of the ORFs encoded in the phage genomes have unknown

functions, and their potential roles in *S. aureus* antibiotic resistance and virulence are unknown, what's more, with the development of genome sequencing, acquisition of a complete genome sequence of organism become much more easily, and lots of ORFs encoded proteins with unknown functions<sup>32</sup>. All of the above factors represent a potential genome-level risk for phage therapy. Therefore, the clinical use of these phages should be evaluated comprehensively.

In summary, *Staphylococcus* phages of the *Myoviridae* family have been isolated worldwide. Morphology studies have shown that these phages are contractile between the head and tail. In addition, they have a wide host range and can prevent infectious diseases in animal models, and pre-clinical and clinical trials have not reported any adverse effects. Our results show that *Staphylococcus* phages of the *Myoviridae* family are lytic and therefore cannot integrate into the genomes of their host. In addition, no antibiotic resistance or virulence genes were present in their genomes, indicating that they are safe at the genome level. In addition, the phages analyzed in this study clustered in the same branch of an evolutionary tree, suggesting that they share a common ancestor.

## References

- Wright, A., Hawkins, C. H., Anggard, E. E. & Harper, D. R. A controlled clinical trial of a therapeutic bacteriophage preparation in chronic otitis due to antibiotic-resistant *Pseudomonas aeruginosa*; a preliminary report of efficacy. *Clin Otolaryngol* **34**, 349–357, doi: COA1973.
- Bruttin, A. & Brussow, H. Human volunteers receiving *Escherichia coli* phage T4 orally: a safety test of phage therapy. *Antimicrob Agents Chemother* **49**, 2874–2878, doi: 49/7/2874.
- Rhoads, D. D. *et al.* Bacteriophage therapy of venous leg ulcers in humans: results of a phase I safety trial. *J Wound Care* **18**, 237–238, 240–233, doi: 10.12968/jowc.2009.18.6.42801 (2009).
- Chen, J. & Novick, R. P. Phage-mediated intergeneric transfer of toxin genes. *Science* **323**, 139–141, doi: 323/5910/139.
- Granobles Velandia, C. V., Kruger, A., Parma, Y. R., Parma, A. E. & Lucchesi, P. M. Differences in Shiga toxin and phage production among stx(2g)-positive STEC strains. *Front Cell Infect Microbiol* **2**, 82, doi: 10.3389/fcimb.2012.00082 (2012).
- Colomer-Lluch, M., Imamovic, L., Jofre, J. & Muniesa, M. Bacteriophages carrying antibiotic resistance genes in fecal wastes from cattle, pigs and poultry. *Antimicrob Agents Chemother*, doi: AAC.00535-11.
- Vandersteegen, K. *et al.* Microbiological and molecular assessment of bacteriophage ISP for the control of *Staphylococcus aureus*. *PLoS One* **6**, e24418, doi: 10.1371/journal.pone.0024418PONE-D-11-09572.
- Kim, M. S. & Myung, H. Complete Genome of *Staphylococcus aureus* Phage SA11. *J Virol* **86**, 10232, doi: 86/18/10232.
- Takemura-Uchiyama, I. *et al.* Genomic and phylogenetic traits of *Staphylococcus* phages S25-3 and S25-4 (family Myoviridae, genus Twort-like viruses). *Ann Microbiol (Paris)*, doi: 10.1007/s13213-013-0762-2 (2013).
- O'Flaherty, S. *et al.* Genome of staphylococcal phage K: a new lineage of Myoviridae infecting gram-positive bacteria with a low G + C content. *Journal of bacteriology* **186**, 2862–2871, doi: 10.1128/JB.186.9.2862-2871 (2004).
- Gutierrez, D. *et al.* Two Phages, phiPLA-RODI and phiPLA-C1C, Lyse Mono- and Dual-Species Staphylococcal Biofilms. *Appl Environ Microbiol* **81**, 3336–3348, doi: 10.1128/AEM.03560-14 (2015).
- Lobocka, M. *et al.* Genomics of Staphylococcal Twort-like Phages - Potential Therapeutics of the Post-Antibiotic Era. *Adv Virus Res* **83**, 143–216, doi: B978-0-12-394438-2.00005-0.
- Cui, Z. *et al.* Complete Genome Sequence of Wide-Host-Range *Staphylococcus aureus* Phage JD007. *J Virol* **86**, 13880–13881, doi: 10.1128/JVI.02728-12 (2012).
- Gu, J. *et al.* Complete Genome Sequence of *Staphylococcus aureus* Bacteriophage GH15. *J Virol* **86**, 8914–8915, doi: 86/16/8914.
- Vandersteegen, K. *et al.* Romulus and Remus, two phage isolates representing a distinct clade within the Twortlikevirus genus, display suitable properties for phage therapy applications. *J Virol* **87**, 3237–3247, doi: 10.1128/JVI.02763-12 (2013).
- Takemura-Uchiyama, I. *et al.* Evaluating efficacy of bacteriophage therapy against *Staphylococcus aureus* infections using a silkworm larval infection model. *FEMS Microbiol Lett.*, doi: 10.1111/1574-6968.12220 (2013).
- McDowall, J. & Hunter, S. InterPro protein classification. *Methods Mol Biol* **694**, 37–47, doi: 10.1007/978-1-60761-977-2\_3 (2011).
- Marchler-Bauer, A. *et al.* CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* **39**, D225–229, doi: 10.1093/nar/gkq1189.
- Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution* **28**, 2731–2739, doi: 10.1093/molbev/msr121 (2011).
- Rissman, A. I. *et al.* Reordering contigs of draft genomes using the Mauve aligner. *Bioinformatics* **25**, 2071–2073, doi: 10.1093/bioinformatics/btp356 (2009).
- McNair, K., Bailey, B. A. & Edwards, R. A. PHACTS, a computational approach to classifying the lifestyle of phages. *Bioinformatics* **28**, 614–618, doi: 10.1093/bioinformatics/bts014 (2012).
- Liu, B. & Pop, M. ARDB—Antibiotic Resistance Genes Database. *Nucleic Acids Res* **37**, D443–447, doi: 10.1093/nar/gkn656 (2009).
- Zhou, C. E. *et al.* MvirDB—a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. *Nucleic Acids Res* **35**, D391–394, doi: 10.1093/nar/gkl791 (2007).
- Drummelsmith, J., Winstall, E., Bergeron, M. G., Poirier, G. G. & Ouellette, M. Comparative proteomics analyses reveal a potential biomarker for the detection of vancomycin-intermediate *Staphylococcus aureus* strains. *Journal of proteome research* **6**, 4690–4702, doi: 10.1021/pr070521m (2007).
- El-Arabi, T. F. *et al.* Genome sequence and analysis of a broad-host range lytic bacteriophage that infects the *Bacillus cereus* group. *Virol J* **10**, 48, doi: 10.1186/1743-422X-10-48 (2013).
- Matos, R. C. *et al.* Enterococcus faecalis prophage dynamics and contributions to pathogenic traits. *PLoS genetics* **9**, e1003539, doi: 10.1371/journal.pgen.1003539 (2013).
- Quiros, P. *et al.* Antibiotic resistance genes in the bacteriophage DNA fraction of human fecal samples. *Antimicrob Agents Chemother* **58**, 606–609, doi: 10.1128/AAC.01684-13 (2014).
- Martinez-Castillo, A., Quiros, P., Navarro, F., Miro, E. & Muniesa, M. Shiga toxin 2-encoding bacteriophages in human fecal samples from healthy individuals. *Appl Environ Microbiol* **79**, 4862–4868, doi: 10.1128/AEM.01158-13 (2013).
- McCallin, S. *et al.* Safety analysis of a Russian phage cocktail: from metagenomic analysis to oral application in healthy human subjects. *Virology* **443**, 187–196, doi: 10.1016/j.virol.2013.05.022 (2013).
- Yang, P., Wang, J. & Qi, Q. Prophage recombinases-mediated genome engineering in *Lactobacillus plantarum*. *Microb Cell Fact* **14**, 154, doi: 10.1186/s12934-015-0344-z (2015).
- De Paepe, M. *et al.* Temperate phages acquire DNA from defective prophages by relaxed homologous recombination: the role of Rad52-like recombinases. *PLoS Genet* **10**, e1004181, doi: 10.1371/journal.pgen.1004181 (2014).
- Hoai, T. D., Nishiki, I. & Yoshida, T. Properties and genomic analysis of *Lactococcus garvieae* lysogenic bacteriophage PLGT-1, a new member of Siphoviridae, with homology to *Lactococcus lactis* phages. *Virus Res* **222**, 13–23, doi: 10.1016/j.virusres.2016.05.021 (2016).
- Kwan, T., Liu, J., DuBow, M., Gros, P. & Pelletier, J. The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. *Proc Natl Acad Sci USA* **102**, 5174–5179, doi: 10.1073/pnas.0501140102.

34. Kvachadze, L. *et al.* Evaluation of lytic activity of staphylococcal bacteriophage Sb-1 against freshly isolated clinical pathogens. *Microbial biotechnology* **4**, 643–650, doi: 10.1111/j.1751-7915.2011.00259.x (2011).
35. 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).
36. El Haddad, L. *et al.* Improving the safety of *Staphylococcus aureus* polyvalent phages by their production on a *Staphylococcus xylosus* strain. *PLoS One* **9**, e102600, doi: 10.1371/journal.pone.0102600 (2014).
37. Alves, D. R. *et al.* Combined use of bacteriophage K and a novel bacteriophage to reduce *Staphylococcus aureus* biofilm formation. *Appl Environ Microbiol* **80**, 6694–6703, doi: 10.1128/AEM.01789-14 (2014).

## Acknowledgements

This work was sponsored in part by the National Natural Science Foundation of China (No. 31500154), the Shanghai Sailing Program (No. 15YF1409500), the fund of the Shanghai Health and Family Planning Committee (No. 201440289), the Pre-outstanding Medical Youth Program of Shanghai General Hospital (No. 06N1503012), and the Shanghai Jiao Tong University School of Medicine Dr Innovation Fund (No. BXJ201201).

## Author Contributions

Z.L.C. and X.K.G. performed the molecular genetic studies, participated in the sequence alignment, and drafted the manuscript. Q.T.L., K.D., L.B.Z., Y.Z., R.T. and Y.Z.Z. participated in the sequence alignment. Z.L.C. and L.L. participated in the design of the study and performed the statistical analysis. Z.L.C. and L.L. conceived of the study, participated in the study design and coordination, and helped to draft the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Cui, Z. *et al.* Safety assessment of *Staphylococcus* phages of the family *Myoviridae* based on complete genome sequences. *Sci. Rep.* **7**, 41259; doi: 10.1038/srep41259 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017