Published in final edited form as:

Gene. 2015 April 25; 561(1): 45–53. doi:10.1016/j.gene.2015.02.053.

Genomic and proteomic features of mycobacteriophage SWU1 isolated from China soil

Xiangyu Fan^{a,b}, Jianlong Yan^a, Longxiang Xie^a, Lanying Zeng^c, Ryland F. Young III^c, and Jianping Xie^{a,*}

Xiangyu Fan: fxysnd@126.com; Jianlong Yan: 519465867@qq.com; Longxiang Xie: xielongxiang123@126.com; Lanying Zeng: lzeng@tamu.edu; Ryland F. Young: ryland@tamu.edu

^aInstitute of Modern Biopharmaceuticals, State Key Laboratory Breeding Base of Eco-Environment and Bio-Resource of the Three Gorges Area, Key Laboratory of Eco-environment of Three Gorges Reservoir, Ministry of Education, School of Life Sciences, Southwest University, Chongqing 400715, China

^bSchool of Biological Science and Technology, University of Jinan, Shandong 250022, China

^cCenter for Phage Technology, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843-2128, USA

Abstract

Mycobacteriophage SWU1 is a newly isolated phage from soil sample collected in Sichuan province, China using *Mycobacterium smegmatis* mc²155 as host. Plaque, phage morphology and one-step growth curve were characterized. The complete genomic sequence of phage SWU1 was determined by shotgun sequencing. The ends of SWU1 were determined. Structural proteins of SWU1 were analyzed by NanoLC-ESI-MS/MS. Seven ORFs were identified as structural protein encoded by SWU1 genome. The genetic basis underlying the SWU1 plaque was explored using comparative genomics. Prophages homologous to SWU1 were identified in two pathogens, *Segniliparus rugosus* ATCC BAA-974 and *Mycobacterium rhodesiae* JS60. Genus *Segniliparus* is a member of the order Corynebacteriales. To our knowledge, this is the first report of *Mycobacterium* prophages in different genera.

Keywords

Mycobacteriophage; Plaque; Electron microscopy; Genome; Proteome

1. Introduction

Bacteriophages represent unprecedented rich bio-resources, especially as sources for novel genetic tools and leads or scaffolds for better antibiotics. Phage-based research tools such as integrative vector and phage display have contributed tremendously to biomedicine and biotechnology. This holds true for mycobacteriophages, a versatile toolkit for *Mycobacterium tuberculosis* molecular manipulation and tuberculosis control (Lee et al.,

*Corresponding author. georgex@swu.edu.cn (J. Xie).

1991; Van Kessel and Hatfull, 2006; Kumar et al., 2008; Marinelli et al., 2008; Van Kessel et al., 2008; Piuri et al., 2009), such as the fluoromycobacteriophages for rapid, specific, and sensitive *M. tuberculosis* antibiotic susceptibility test, and more recently, the mycobacterial recombineering system enabling more simpler and more efficient gene mutagenesis in *Mycobacteria*.

Since the first successful isolation of mycobacteriophage in 1947 (McNerney, 1999), sequenced mycobacteriophages in GeneBank alone have exceeded 600 (Pedulla et al., 2003; Hatfull et al., 2006, 2008; Hatfull, 2010; Hatfull et al., 2010; Pope et al., 2011; Hatfull, 2012) (<http://phagesdb.org/>). However, they are just the tips of the immense mycobacteriophage iceberg (Hatfull et al., 2010; Pope et al., 2011). Given that currently available mycobacteriophages are largely isolated from North America, and many prophages and the CRISPR (clustered regularly interspaced short palindromic repeats) spacers within mycobacterial genomes failed to find matching phages (Hatfull, 2010), more mycobacteriophages should be isolated if samples were more diverse, and this could further enrich our appreciation of this highly diverse and important phage family.

In this paper, a novel mycobacteriophage (SWU1) isolated from China soil, its plaque morphology, electron micrographs, one-step growth curve, genome features and structural proteins were described. Genetic basis underlying this unique plaque was clarified using comparative genomics. In addition, an unexpected homologous SWU1-like prophage was found from *Segniliparus rugosus* ATCC BAA-974 and *Mycobacterium rhodesiae* JS60.

2. Materials and methods

2.1. Specimens and ethics statement

Moist soil samples near a pigpen were collected and preserved in sterile containers from the Gongping village, Sima Town, Pingchang County, Sichuan Province, China (31°45'N, 107°6'E). No specific permits were required for the described field studies in China. The location is not privately-owned. To our knowledge, the field studies do not involve endangered or protected species. *Mycobacterium smegmatis* mc²155 was chosen as viral host for its non-pathogenic nature and shorter generation time.

2.2. Isolation of phages

About 10 g of the soil sample, 15 ml of MP buffer (Tris-HCl pH 7.5 50 mM, NaCl 150 mM, MgCl₂ 10 mM, CaCl₂ 2 mM), 30 µl of AMP (50 mg/ml) and 1 ml *M. smegmatis* mc²155 suspension (without Tween; OD₆₀₀ = 0.45) were added and incubated at 37 °C overnight. Remove insoluble debris by centrifugation at 5000 *g* for 10 min at 4 °C. The supernatant was filtered through a 0.22 µm filtration membrane. About 2 ml of the filtrate and 1 ml *M. smegmatis* mc²155 suspension (without tween; OD₆₀₀ = 0.45) were mixed and incubated at 37 °C for 30 min. About 1 ml of the cell-phage mixture and 3 ml of soft agar (0.7% agar, 1 mM CaCl₂) were mixed and poured on 7H10/AMP plates. The plates were incubated at 30 °C for 2–3 days. Typical single plaque was picked for at least 5 generations until a purified phage is obtained and high titer lysates were prepared.

2.3. Phage purification and DNA isolation

20 µl of phage lysates was mixed with 1 ml of late-log phase *M. smegmatis* mc²155 cells and 10 ml of soft agar (0.7% agar, 1 mM CaCl₂). This mixture was poured on a 150 mm plate containing 7H9 agar (7H9 broth base + 2% agar). Eighteen such plates were prepared. Allow plates to harden, and then incubate at 30 °C overnight. Add 10 ml MP buffer to each plate. Following overnight incubation at 4 °C, the buffer was pipetted off the plate. Centrifuge 3500 *g* for 5 min at 4 °C to pellet bacterial cells. Transfer the supernatant to a 250 ml flat bottom plastic bottle. Add 10% polyethylene glycol (PEG 4000) and NaCl to 1 M. Stir this at 4 °C overnight. Centrifuge the precipitated phage for 5 min 3500 *g* at 4 °C. Resuspend the phage pellets in 6 ml MP buffer. Pipette gently up and down, rock at 4 °C overnight. Centrifuge at 3500 *g* at 40 °C, and then save the clear supernatant. The phage was purified using CsCl density gradient centrifugation with *d* = 1.45 upper and *d* = 1.50 lower CsCl layers. The phage band at the interface between the two CsCl layers was removed with a syringe and needle. Dialyze repeatedly at 4 °C in MP buffer to remove the CsCl. Use phenol to break up the phage particles and isolate the DNA. The DNA was ethanol precipitated and resuspended in ddH₂O.

2.4. Characterization of phage biology

2.4.1. Plaque morphology—Separated and two adjacent plaques were chosen for further morphology characterization. The parameters for clear/turbid nature and appearance of the plaques were noted. 500 µl of MP buffer, 500 µl of *M. smegmatis* mc²155 suspension (without Tween; OD₆₀₀ = 0.45) and 4.5 ml of soft agar (0.7% agar, 1 mM CaCl₂) were mixed and poured on a 100 mm plate containing 7H10 agar. Spot 5 µl SWU1 phage lysate (4×10^5 PFU) on the plate after the top agar has cooled and hardened completely. After the spots evaporate, incubate at 37 °C overnight. The size of the halo surrounding the plaques was measured.

2.4.2. Phage morphology—Purified SWU1 phage lysate was stained with 2% phosphotungstic acid. And then electron micrographs of SWU1 were observed under electron microscope (FEI TECNAI10). The electron microscopy was manipulated by a technician in Third Military Medical University.

2.5. One-step growth curve

A standardized protocol (Huang et al., 2013) was adapted for the mycobacteriophages. A log-phase culture (200 ml) of *M. smegmatis* mc²155 (OD₆₀₀ = 0.9) was harvested by centrifugation and washed twice to remove Tween-80 using fresh 7H9 (without Tween-80). Mycobacteriophage SWU1 was added and allowed to absorb for 30 min at 37 °C using a MOI of 10. The mixture was then centrifuged at 12,000 rpm for 1 min. Pelleted cells were washed twice to remove unabsorbed phage and resuspended using MP buffer. Cells were incubated at 37 °C, and samples were taken every 45 min to monitor the titer.

2.6. DNA preparation and sequencing

The SWU1 DNA (concentration 50 ng/µl, total amount of DNA 20 µg, OD_{260/280}:1.80–2.00) was prepared to construct chromosomal DNA library. 20 µg phage

DNA was sheared with HydroShear DNA Fragmenting System (Gene Machines, #HSA-01-X-00109). 80 µl DNA fragment, 5 µl T4 DNA Polymerase (TaKaRa, 5 U/µl), 10 µl 10 × T4 DNA polymerase buffer and 5 µl dNTP (10 mM) were added to a 0.5 ml EP tube and incubated for 1 h at 37 °C by a water bath. The phage DNA fragments were size-fractionated on a 0.7% agarose gel. The 1.5–3 kb DNA size fraction was electroeluted from the gel and the corresponding fragments cloned into PUC118. The resulting ligation mixture was used to transform *Escherichia coli* DH10B by electroporation, and insert-containing white colonies were grown overnight in LB medium for plasmid extraction. The extracted plasmid was sequenced from both ends on the Applied Biosystem ABI PRISM™3730 DNA Sequencer. Random phage sequences were assembled and edited using Sequencer software (Phred/Phrap/Consed). The sequence was completed using a small number of primers constructed to fill in gaps and either a double stranded clone or whole SWU1 genome as template DNA. The service of sequencing was commercially completed in BGI in China.

2.7. Identification of genome ends

The cohesive termini of SWU1 genome were sequenced using PCR primers (A: 5'-TGGCACAGAAGGCCGTCG C-3', and B: 5'-GCGTCATCCATCGCCACCTTGATGG-3') extending outward from each end of the genome using whole SWU1 genomic DNA as template.

2.8. Sequence analysis

The nucleotide composition of SWU1 was analyzed using Geneious 5.3.6 downloaded at <http://www.geneious.com/default,83,home.sm>. Multiple sequences alignment between SWU1 and L5 was analyzed using DNAMAN. Predicted genes were subjected to database searches using the Glimmer (Delcher et al., 2007). The presence of putative tRNA and tmRNA was screened using tRNAscan-SE Search Server (<http://lowelab.ucsc.edu/tRNAscan-SE/>) (Schattner et al., 2005). Promoter prediction was identified by using the Berkeley Drosophila Genome Project (BDGP) available at http://www.fruitfly.org/seq_tools/promoter.html (NNPP). Terminator prediction was identified using FindTerm program available at <http://molbiol-tools.ca/Promoters.htm>. The use of a previously constructed L5 codon usage table assisted the identification of SWU1 predicted genes (Hatfull and Sarkis, 1993).

2.9. Analysis of SWU1 structural proteins

The purified SWU1 was resuspended in 36 µl 5 × SDS-PAGE loading buffer and 4 µl β-mercaptoethanol. The sample was boiled for 30 min and 10 µl of it was loaded onto a 15% (w/v) polyacrylamide gel. The gel that contained structural proteins of phage SWU1 was cut and digested for NanoLC-ESI-MS/MS analysis in ProfTech Inc.

3. Results

3.1. Biological characteristics of SWU1

Using *M. smegmatis* mc²155 as host, one mycobacteriophage was isolated from the soil sample. This phage was named SWU1. Plaques of SWU1 appeared clear, without the halo on *M. smegmatis* mc²155 lawn after overnight incubation. From the next day onward, the

plaques showed four unique circles which alternated as clear, turbid, clear, and turbid outwards. Once formed, the halo remained the same size throughout the culture duration. The typical plaque pattern disappeared when one clone hugs another, and replaced by a straight line between the two clones (Fig. 1A).

SWU1 showed a normal morphology under a transmission electron microscope (Fig. 1B). The phage head is isometric dodecahedron while the tail is long.

All parameters of the lytic cycle of phage SWU1 were determined from a one-step growth curve. The latent and burst time are calculated to be 30 and 270 min, respectively (Fig. 2). The average burst size is estimated to be about 200 PFU per infected host bacterium.

3.2. Overview of SWU1 genome

Genome announcement of SWU1 had been published (Fan et al., 2012). The GenBank accession number for SWU1 is JF946695. Bioinformatic analysis reveals that SWU1 and mycobacteriophage L5 attachment sites (*attP*) are very alike based on the genome sequence, although two positions at the left end are different. The SWU1 *attP* (coordinate of the common core range from 25,492 to 25,534) locates somewhat to the right center of the genome and divides the genome into a left and a right arm (Fig. 3). The SWU1 genome was annotated to have 94 putative protein-coding genes, 3 tRNA genes and without small RNA genes. Of these 94 genes, 30 ORFs locate in the left arm and 64 in the right arm (Fig. 3). And 91 of these putative genes have identifiable sequence similarity to other mycobacteriophage genes at the amino acid sequence level and 22 genes have been assigned functions (Additional file 1), although at least three fail to match with any other genes.

The majority of genes in the left arm are transcribed rightward, and only gene 65, the phage integrase, is transcribed leftward. The majority of genes in the right arm are transcribed leftward and only gene 64 and gene 63 are transcribed rightward. Overall, noncoding spaces are rare, with the notable exceptions of ~450 bp between the left end of the genome and gene 94, ~250 bp between genes 65 and *attP* site, ~300 bp between genes 20 and 19, and ~1200 bp between gene 1 and the right end of the genome. Applying a time-delay neural network for promoter annotation, we have identified two putative promoters in the SWU1 genome. One promoter locates on the noncoding spaces between gene 64 and *attP* site and links to gene 64. Another locates on the noncoding spaces between gene 20 and gene 19 and links to leftward-transcribed genes (Fig. 3). There are five putative terminators in the SWU1 genome using FindTerm program terminator discovery. As described in Fig. 3, all terminators locate on the noncoding spaces among genes and link to operons.

Many of the SWU1 left arm genes are involved in phage structure and assembly. 13 of 33 genes are clearly related to proteins with established functions in other mycobacteriophages. Genes 85, 84, 82 and 81 encode terminase, portal protein, scaffold protein and major head subunit, respectively. Gene 75 encodes the major tail subunit. Genes 89, 72, 71 and 70 all encode the minor tail subunit, and gp72 may be a tape measure protein. Genes 88, 87 and 86 encode LysA, Holin and LysB, which are components of the lytic cassette of mycobacteriophage SWU1. Gene 65 encodes integrase highly similar to that of L5 (99% identity, E-value, 0).

Several genes in the right arm of the SWU1 genome appear to encode proteins that are involved in phage DNA metabolism, based on sequence similarity to proteins with established functions of other mycobacteriophages. Genes 54, 49 and 48 encode DNA polymerase, flavin-dependent thymidylate synthase (ThyX) and class II ribonucleotide reductase (RNR), respectively. Gene 62 encodes phage excisionase. The phage integrase (gp65), the phage attachment site (25,492–25,534) and the excisionase (gp62) form an integration cassette, suggesting that SWU1 is capable of developing lysogen or possibly lost its lysogenic functions during a lengthy natural selection. Genes 14, 13 and 12 are highly similar to the 77, 78 and 79 of mycobacteriophage L5. They are three cytotoxic early proteins and can inhibit the growth of host *M. smegmatis* (Rybniiker et al., 2008).

We note that gp51 is significantly similar (87% identity, E-value, 9e-63) to a putative DNA binding protein of mycobacteriophage L5. The protein contains an HTH_Hin_like superfamily DNA binding domain, suggesting an ssDNA annealing protein (SSAP). Gp35 is obviously similar (63% identity, E-value, 9e-36) to recombination endonuclease VII of *S. rugosus* ATCC BAA-974. Gp23 is similar to RecB family exonuclease of *Actinomyces* sp. oral taxon 178 str. F0338 (27% identity, E-value, 9e-11). Based on the above analysis we conclude that the SWU1 genome contains functional genes of homologous recombination. Endonuclease specifically binds and cleaves Holliday junction to yield nicked double-stranded DNA (dsDNA). Exonuclease degrades a single strand of the nicked dsDNA, thereby exposing a 3' single-stranded DNA (ssDNA) tail. The SSAP binds 3' ssDNA ends, forming a protein-DNA filament, which protects the substrate DNA and promotes the annealing of complementary DNA strands, strand exchange and strand invasion. However, endonuclease, exonuclease and SSAP of SWU1 locate at non-adjacent locus. Therefore, it remains an open question if they are bound together to play a role in the homologous recombination.

3.3. Identification of the genome end

Two PCR primers, namely A and B outward from each end of the genome, and the whole SWU1 genome DNA as template (Fig. 4A) were used to amplify the ends and the products have been successfully obtained (Fig. 4B), suggesting the existence of cohesive termini within SWU1 genome and formed a continuous template during PCR amplification. The PCR product should contain the cohesive termini.

The PCR product and each end of the SWU1 genome were sequenced by PCR primers A and B. The results show that the SWU1 genome is 5'-CGGTCGGTTA-3' with a 3' extension (Fig. 4C).

3.4. Genetic mosaicism of SWU1 and the genetic basis of novel plaque

We noted that SWU1 has 94.66% identity to mycobacteriophage L5 by multiple sequence alignments in DNAMAN. And in NCBI 'megablast' comparison, the extents of the similarity cover 97% of the genomes, and all segments have greater than 88% identity. Although the DNA sequences of SWU1 and L5 are pretty alike, their morphologies are remarkably different. L5 forms turbid plaques, while SWU1 exhibits a characteristic alternative clear-turbid-clear-turbid circle on *M. smegmatis* mc² 155 lawns. The observed slight genome

difference might be responsible for the observed plaque appearance. However, further mutagenesis study is needed to confirm this hypothesis. Similar to other mycobacteriophages, SWU1 genome is mosaic. As described in Fig. 5A, two segments of genomic differences are present. There are many insertions, deletions, and substitutions.

3.4.1. Insertions and deletions—There are two examples where SWU1 contains DNA segments not present in L5. As described in Fig. 5B, SWU1 genes 58 and 56 have significant similarities to the genes 41 and 42 of L5, but SWU1 has a gene that L5 lacks: gene 57. However, the gene is similar to Mycobacteriophage Pukovnik gene 41 and has weak sequence similarity to acyl-CoA carboxylase carboxyl transferase beta subunit of *Gramella forsetii* KT0803, suggesting a role in lipid metabolism. In another example, SWU1 has a DNA segment, which contains a newly assigned gene SWU1 gene 39, inserted between genes 40 and 38. And SWU1 genes 40 and 38 have significant similarities to genes 56 and 57 of L5 (Fig. 5C). SWU1 gene 39 has weak sequence similarity to leishmanolysin-like peptidase of *Clonorchis sinensis* (35% identity, E-value, 0.54) and putative head-binding domain of phage tailspike protein of *Escherichia* sp. 4_1_40B (41% identity, E-value, 6.6). The tailspike protein can recognize and bind to the receptor of the host cell. So SWU1 gp39 may be a depolymerase. The purified recombinant SWU1 gp39 does have depolymerase activity (our unpublished observations).

As described in Fig. 5B, the L5 gene 40 is absent in SWU1. Little is known about the function of L5 gp40. It is worthwhile to explore whether the deletion of the L5 gene 40 affects SWU1 particles.

3.4.2. Gene substitutions—A case of gene substitution is apparent, in which an established L5 gene 45 is absent in SWU1, replaced by novel SWU1 genes 53 and 52 (Fig. 5B). SWU1 gp53 has weak sequence similarity to hypothetical protein MycrhDRAFT_5777 of *M. rhodesiae* JS60 (41% identity, E-value, 5e-04), transcriptional regulator LacI family of *Jonesia denitrificans* DSM 20603 (41% identity, E-value, 0.89) and gp47 of Mycobacteriophage RedRock (36% identity, E-value, 9.9), suggesting a role in host transcriptional response. SWU1 gp52 is a close relative of D29gp43 (60% identity, E-value, 4e-44) and has weak sequence similarity to major head protein (gp17) of L5 (50% identity, E-value, 2.7), but only has 26% coverage.

3.5. Identification of SWU1 structural proteins

The sodium dodecyl sulfate-polyacryl-amide gel that contained structural proteins of SWU1 was cut, digested and subjected to NanoLC-ESI-MS/MS analysis. Seven structural proteins were identified. Six of them were located on the left arm, and one of them (gp42) was located on the right arm (Table 1). However, some predicted structural proteins such as gp89, gp82 and gp70 failed to be identified.

3.6. SWU1-like prophage

To our surprise, database searches reveal that 32 genes in SWU1 can match to the genome of *S. rugosus* ATCC BAA-974 (Table 2) and 54 genes are closest relative to the genome of *M. rhodesiae* JS60 (Table 3), suggesting a prophage in *S. rugosus* ATCC BAA-974 and *M.*

rhodesiae JS60 homologous to SWU1. We cannot get complete genome sequences of these two bacteria, so the sequences of the two prophages cannot be downloaded for further analysis. In the two prophages, all constitutive proteins, as well as some proteins involved in phage nucleic acid metabolism of SWU1 can be found. However, no integrase and excisionase involving in the switch of phage from lysogenic state into lytic state can be found in prophages. *S. rugosus* ATCC BAA-974 (Butler et al., 2005, 2007; Hansen et al., 2009) is an opportunistic pathogen, belonging to the *Actinobacteria Actinomycetales Corynebacterineae*. *M. rhodesiae*, a rapidly growing scotochromogenic mycobacterium, was first isolated from the sputum of pulmonary tuberculosis patients in Zimbabwe in 1971 (Tsukamura et al., 1971). And it was reported in a case of continuous ambulatory peritoneal dialysis (CAPD)-associated peritonitis due to *M. rhodesiae* (Curry et al., 2008). At first glance, the two bacteria are distant. However, the cell walls of both *S. rugosus* and *Mycobacterium* are rich in mycolic acids. *S. rugosus* has ultralong C100 mycolic acids (Hong et al., 2012). *S. rugosus* is an emerging important human pathogen (Kim et al., 2013) isolated from the respiratory tract of human and non-human mammals. It is no wonder they share a homologous prophage. Many bacterial virulence genes are located within prophages of the bacteria (Wagner and Waldor, 2002). But there is no evidence that these prophages contribute to virulence. However, this phenomenon, to our knowledge, is firstly reported.

4. Discussion

In conclusion, we present here the biology and genomic and proteomic characteristics of mycobacteriophages SWU1 isolated from soil sample in China. Plaque morphology of the temperate mycobacteriophage varies significantly (Kumar et al., 2008). The mycobacteriophage SWU1 exhibits characteristic alternative clear–turbid–clear–turbid circles on *M. smegmatis* mc²155 lawn. The plaque feature is the opaque halo surrounding the clear lysis zone. Bacteriophages associated with host capsule depolymerases usually form plaques with halos (Rieger-Hug and Stirm, 1981) due to the phage released depolymerization enzyme capable of hydrolyzing the receptor polysaccharide and the macromolecular structure of the bacterial capsule (Bayer et al., 1979). More often than not, these depolymerases are associated with a component of the phage tails and also can degrade biofilm. Therefore, we speculated that mycobacteriophage SWU1 genome might harbor such depolymerases, which can explain the plaque difference between SWU1 and L5. Two segments of genomic differences are present. In conclusion, SWU1 gp57, 53, 52 and 39 are absent in L5; L5 gp40 and gp45 are absent in SWU1. These proteins might hold the key to unveil the microevolution of SWU1 which produces novel plaque morphology. A BLASTP search revealed that amino acid sequences of SWU1 gp57, 53 and 52 are highly similar to proteins in other mycobacteriophages. These mycobacteriophages lack the morphology of SWU1 on *M. smegmatis* mc² 155 lawns, which preclude the role of SWU1 gp57, 53 and 52 in the formation of SWU1 plaque. Our preliminary analysis suggests that SWU1 gp39 might be related to the observed novel plaque of SWU1.

It seems that the lines between the two plaques are lawns formed by phage resistant bacteria. Mycobacteriophage SWU1 cannot infect these putative resistant bacteria. However, more works are needed to elucidate the underlying mechanisms.

We have identified structural proteins of SWU1. We are very interested with one of those structural proteins, gp67. The protein belongs to pham 7406 (<http://phagesdb.org/>). Some proteins (Halo gp22, Rosebush gp32) of this pham are proved to be involved in host range (Jacobs-Sera et al., 2012). SWU1 gp67 only accounts for 0.4% of structural proteins. SWU1 gp67 might be a tail fiber protein involved in adsorbing to *Mycobacterium*.

We confirmed that the genome sequences of *S. rugosus* ATCC BAA-974 and *M. rhodesiae* JS60 prophage are similar to SWU1 through BLAST-P. This implies that some phage can transmit cross-genus. *M. smegmatis* was confirmed to be able to support the propagation of DNA of *Streptomyces* phage Zemlya Jacobs-Sera et al. (2012). Mycobacteriophage subclusters A1 and A2 have sequence similarity to R4 cluster *Streptomyces* phages (Smith et al., 2013). *S. rugosus* is just like *Streptomyces* that is closely related to *Mycobacterium*. However, whether *S. rugosus* can be infected by SWU1 or other mycobacteriophages remains open. It is possible that the ancestor of *S. rugosus* infected with SWU1 became lysogen, followed by the loss of genes required for the lysis of phage during a lengthy natural evolution. Thereby deposited the prophage in the genome of *S. rugosus* is a testimony of evolution. This might be an example of bacteriophage transmission between different bacteria. However, how the remnants of the SWU1 phage can be found in two different genera remains enigmatic. Further genome sequence scrutiny might reveal the evolution of the two species. Many bacterial virulence genes are located within prophages of the bacteria (Wagner and Waldor, 2002). But it is unclear whether these prophages contribute to the pathogenesis of *S. rugosus* and *M. rhodesiae*. However, this phenomenon, to our knowledge, is firstly reported.

5. Conclusions

In conclusion, we presented here a novel mycobacteriophage SWU1, which was isolated from a soil sample in China. The analyses of morphology, biological characteristics, and genomic characteristics were performed. We are interested in several features of SWU1, e.g. the plaque pattern and underlying genetic basis. SWU1 gp39 was thought to be related to the observed novel plaque of SWU1 using comparative genomics. Structural proteins of SWU1 also were identified using NanoLC-ESI-MS/MS analysis. We also confirmed that the genome sequences of *S. rugosus* ATCC BAA-974 and *M. rhodesiae* JS60 prophage were similar to SWU1 through Blast-P.

Acknowledgments

This work was supported by the National Natural Science Foundation (grant numbers 81371851, 81071316, 81271882, 81301394); the New Century Excellent Talents in Universities (grant number NCET-11-0703); the National Megaprojects for Key Infectious Diseases (grant number 2008ZX10003-006); the Excellent PhD Thesis Fellowship of Southwest University (grant numbers kb2010017, ky2011003); the Fundamental Research Funds for the Central Universities (grant numbers XDJK2011D006, XDJK2012D011, XDJK2012D007, XDJK2013D003 and XDJK2014D040); the Postgraduates Innovative Project of Chongqing Municipal Council of Education (grant number CYS14044); the Natural Science Foundation Project of CQ CSTC (grant number CSTC 2010BB5002); the Chongqing Municipal Committee of Education for Postgraduates Excellence Program (grant number YJG123104); the Undergraduates Teaching Reform Program (grant numbers 2011JY052, 2013JY201); the Shandong Excellent Young Scientist Award Fund (grant number BS2014YY031).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.gene.2015.02.053>.

Abbreviations

AMP	ampicillin
BLAST	Basic Local Alignment Search Tool
bp	base pair
CRISPR	clustered regularly interspaced short palindromic repeats
MOI	multiplicity of infection
NCBI	National Center for Biotechnology Information
ORF	open reading frame
PCR	polymerase chain reaction

References

- Bayer ME, Thurow H, Bayer MH. Penetration of the polysaccharide capsule of *Escherichia coli* (B161/42) by bacteriophage K29. *Virology*. 1979; 94:95–118. [PubMed: 375578]
- Butler W, Floyd MM, Brown JM, Toney SR, Daneshvar MI, Cooksey RC, Carr J, Steigerwalt AG, Charles N. Novel mycolic acid-containing bacteria in the family *Segniliparaceae* fam.nov., including the genus *Segniliparus* gen. nov., with descriptions of *Segniliparus rotundus* sp.nov and *Segniliparus rugosus* sp.nov. *Int J Syst Evol Microbiol*. 2005; 55:1615–1624. [PubMed: 16014491]
- Butler W, Sheils CA, Brown-Elliott BA, Charles N, Colin AA, Gant MJ, Goodill J, Hindman D, Toney SR, Wallace RJ Jr. First isolations of *Segniliparus rugosus* from patients with cystic fibrosis. *J Clin Microbiol*. 2007; 45:3449–3452. [PubMed: 17670929]
- Curry E, Yehia M, Roberts S. CAPD peritonitis caused by *Mycobacterium rhodesiae*. *Perit Dial Int*. 2008; 28:97–99. [PubMed: 18178955]
- Delcher AL, Bratke KA, Powers EC, Salzberg SL. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics*. 2007; 23:673–679. [PubMed: 17237039]
- Fan X, Teng T, Wang H, Xie J. Biology of a novel mycobacteriophage, SWU1, isolated from Chinese soil as revealed by genomic characteristics. *J Virol*. 2012; 86:10230–10231. [PubMed: 22923793]
- Hansen T, Van Kerckhof J, Jelfs P, Wainwright C, Ryan P, Coulter C. *Segniliparus rugosus* infection, Australia. *Emerg Infect Dis*. 2009; 15:611. [PubMed: 19331748]
- Hatfull GF. Mycobacteriophages: genes and genomes. *Annu Rev Microbiol*. 2010; 64:331–356. [PubMed: 20528690]
- Hatfull GF. Complete genome sequences of 138 mycobacteriophages. *J Virol*. 2012; 86:2382–2384. [PubMed: 22282335]
- Hatfull GF, Pedulla ML, Jacobs-Sera D, Cichon PM, Foley A, Ford ME, Gonda RM, Houtz JM, Hryckowian AJ, Kelchner VA. Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genet*. 2006; 2:e92. [PubMed: 16789831]
- Hatfull GF, Cresawn SG, Hendrix RW. Comparative genomics of the mycobacteriophages: insights into bacteriophage evolution. *Res Microbiol*. 2008; 159:332–339. [PubMed: 18653319]
- Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko CC, Weber RJ, Patel MC, Germane KL, Edgar RH. Comparative genomic analysis of 60 mycobacteriophage genomes:

- genome clustering, gene acquisition, and gene size. *J Mol Biol.* 2010; 397:119–143. [PubMed: 20064525]
- Hatfull GF, Sarkis GJ. DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. *Mol Microbiol.* 1993; 7:395–405. [PubMed: 8459766]
- Hong S, Cheng TY, Layre E, Sweet L, Young DC, Posey JE, Butler WR, Moody DB. Ultralong C100 mycolic acids support the assignment of *Segniliparus* as a new bacterial genus. *PLoS One.* 2012; 7:e39017. [PubMed: 22720018]
- Huang G, Le S, Peng Y, Zhao Y, Yin S, Zhang L, Yao X, Tan Y, Li M, Hu F. Characterization and genome sequencing of phage Abp1, a new phiKMV-like virus infecting multidrug-resistant *Acinetobacter baumannii*. *Curr Microbiol.* 2013; 66:535–543. [PubMed: 23328903]
- Jacobs-Sera D, Marinelli LJ, Bowman C, Broussard GW, Guerrero Bustamante C, Boyle MM, Petrova ZO, Dedrick RM, Pope WH, Modlin RL, Hendrix RW, Hatfull GF. On the nature of mycobacteriophage diversity and host preference. *Virology.* 2012; 434:187–201. [PubMed: 23084079]
- Kim JS, Kim WS, Lee K, Won CJ, Kim JM, Eum SY, Koh WJ, Shin SJ. Differential immune responses to *Segniliparus rotundus* and *Segniliparus rugosus* infection and analysis of their comparative virulence profiles. *PLoS One.* 2013; 8:e59646. [PubMed: 23555735]
- Kumar V, Loganathan P, Sivaramakrishnan G, Kriakov J, Dusthakeer A, Subramanyam B, Chan J, Jacobs WR Jr, Paranj Rama N. Characterization of temperate phage Che12 and construction of a new tool for diagnosis of tuberculosis. *Tuberculosis (Edinb).* 2008; 88:616–623. [PubMed: 18511339]
- Lee MH, Pascopella L, Jacobs WR, Hatfull GF. Site-specific integration of mycobacteriophage L5: integration-proficient vectors for *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and bacille Calmette–Guerin. *Proc Natl Acad Sci U S A.* 1991; 88:3111–3115. [PubMed: 1901654]
- Marinelli LJ, Piuri M, Swigonova Z, Balachandran A, Oldfield LM, van Kessel JC, Hatfull GF. BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. *PLoS One.* 2008; 3:e3957. [PubMed: 19088849]
- McNerney R. TB: the return of the phage. A review of fifty years of mycobacteriophage research. *Int J Tuberc Lung Dis.* 1999; 3:179–184. [PubMed: 10094316]
- Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, Lewis JA, Jacobs-Sera D, Falbo J, Gross J, Pannunzio NR. Origins of highly mosaic mycobacteriophage genomes. *Cell.* 2003; 113:171–182. [PubMed: 12705866]
- Piuri M, Jacobs WR Jr, Hatfull GF. Fluoromycobacteriophages for rapid, specific, and sensitive antibiotic susceptibility testing of *Mycobacterium tuberculosis*. *PLoS One.* 2009; 4:e4870. [PubMed: 19300517]
- Pope WH, Jacobs-Sera D, Russell DA, Peebles CL, Al-Atrache Z, Alcoser TA, Alexander LM, Alfano MB, Alford ST, Amy NE. Expanding the diversity of mycobacteriophages: insights into genome architecture and evolution. *PLoS One.* 2011; 6:e16329. [PubMed: 21298013]
- Rieger-Hug D, Stirm S. Comparative study of host capsule depolymerases associated with *Klebsiella* bacteriophages. *Virology.* 1981; 113:363–378. [PubMed: 7269247]
- Rybníček J, Plum G, Robinson N, Small PL, Hartmann P. Identification of three cytotoxic early proteins of mycobacteriophage L5 leading to growth inhibition in *Mycobacterium smegmatis*. *Microbiology.* 2008; 154:2304–2314. [PubMed: 18667563]
- Schattner P, Brooks AN, Lowe TM. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* 2005; 33:W686–W689. [PubMed: 15980563]
- Smith MC, Hendrix RW, Dedrick R, Mitchell K, Ko CC, Russell D, Bell E, Gregory M, Bibb MJ, Pethick F, Jacobs-Sera D, Herron P, Buttner MJ, Hatfull GF. Evolutionary relationships among actinophages and a putative adaptation for growth in *Streptomyces* spp. *J Bacteriol.* 2013; 195:4924–4935. [PubMed: 23995638]
- Tsukamura M, Mizuno S, Gane N, Mills A, King L. *Mycobacterium rhodesiae* sp.nov. A new species of rapid-growing scotochromogenic mycobacteria. *Jpn J Microbiol.* 1971; 15:407–416. [PubMed: 5316571]

- Van Kessel J, Hatfull G. Recombineering in *Mycobacterium tuberculosis*. Nat Methods. 2006; 4:147–152. [PubMed: 17179933]
- Van Kessel J, Marinelli L, Hatfull G. Recombineering mycobacteria and their phages. Nat Rev Microbiol. 2008; 6:851–857. [PubMed: 18923412]
- Wagner PL, Waldor MK. Bacteriophage control of bacterial virulence. Infect Immun. 2002; 70:3985–3993. [PubMed: 12117903]

Fig. 1.

The plaques and morphology of SWU1. A) Plaque of SWU1 presents a characteristic bull's eye when plated it on *M. smegmatis* mc2155 lawn and the juncture of neighboring plaques forms line style lawn. B) Electron micrograph shows that SWU1 has an isometric, icosahedral head and a long flexible tail. Scale bar = 100 nm.

Fig. 2. One-step growth curve of phage SWU1. Experiments were repeated three times with duplicate samples. The Y-axis shows the plaque forming units per milliliter (PFU/ml). The X-axis indicates the time.

Fig. 3.
Organization of the mycobacteriophage SWU1 genome.

Fig. 4.

Identification of SWU1 genome ends. A) PCR diagram: primers A and B are designed close to the two ends of the genome and PCR results in amplified products as shown after agarose gel electrophoresis. B) Lane 1, the PCR product. Lane 2, DL2000 marker, the size of the bands is indicated on the right. C) The cohesive end sequence in bold was identified by comparing the two end fragment sequences with the PCR fragment.

Fig. 5. Global comparison of SWU1 and L5. A) Two segments of genomic differences are present. There are a lot of insertions, deletions, and substitutions. B and C) We note that SWU1 contains genes that L5 does not: gene 57, 53, 52 and 39; L5 gp40 and gp 45 are not present in SWU1.

Table 1

SWU1 virion proteins determined by mass spectrometry.

Virion protein of SWU1	Function	Relative abundance
Gp84	Portal protein	63.3%
Gp81	Major head subunit	13.6%
Gp75	Major tail subunit	10.4%
Gp72	Minor tail subunit (TMP)	8.7%
Gp71	Minor tail subunit	3.3%
Gp67	Virion protein	0.4%
Gp42	Virion protein	0.3%

Table 2

The genes which have homology to *S. rugosus* ATCC BAA-974 in SWU1.

The protein of SWU1	The protein of <i>S. rugosus</i> ATCC BAA-974	Function	Identity
Gp92	Hypothetical protein HMPREF9336_02215		38%
Gp91	Hypothetical protein HMPREF9336_02214	Minor tail protein	42%
Gp88	Hypothetical protein HMPREF9336_02210	Lysin B	63%
Gp87	Hypothetical protein HMPREF9336_02209	Terminase	64%
Gp86	Hypothetical protein HMPREF9336_02208	Portal protein	51%
Gp85	Hypothetical protein HMPREF9336_02207		48%
Gp84	Hypothetical protein HMPREF9336_02206	Head assembly protein	36%
Gp83	Hypothetical protein HMPREF9336_02205	Major head protein	27%
Gp81	Phage protein Gp19/Gp15/Gp42		47%
Gp80	Hypothetical protein HMPREF9336_02203		50%
Gp79	Zinc import ATP-binding protein znuC		32%
Gp78	Hypothetical protein HMPREF9336_02201		64%
Gp77	Major tail protein Gp23	Major tail protein	54%
Gp76	Hypothetical protein HMPREF9336_02199		36%
Gp75	D site-binding protein		45%
Gp74	Cation diffusion facilitator family transporter	Minor tail protein	28%
Gp73	Minor tail protein Gp27	Minor tail protein	59%
Gp72	Minor tail protein Gp28	Minor tail protein	64%
Gp71	Hypothetical protein HMPREF9336_02194		53%
Gp69	Hypothetical protein HMPREF9336_02191		32%
Gp61	Hypothetical protein HMPREF9336_02248		40%
Gp56	DNA polymerase family A	DNA polymerase	56%
Gp53	Chaperone	Probable DNA binding protein	37%
Gp49	Formin protein		25%
Gp46	Hypothetical protein HMPREF9336_02244		43%
Gp44	Hypothetical protein HMPREF9336_02243		29%
Gp43	Phosphoesterase		66%
Gp42	Hypothetical protein HMPREF9336_02239		46%
Gp41	Glutaredoxin		46%
Gp39	Toprim domain-containing protein		55%
Gp37	Recombination endonuclease VII	Endonuclease	63%
Gp25	Hypothetical protein HMPREF9336_02220		51%

Table 3

The genes which have homology to *M. rhodesiae* JS60 in SWU1.

The protein in SWU1	The protein in <i>M. rhodesiae</i> JS60	The predicted function of protein in SWU1	Identity
Gp92	Protein of unknown function DUF124		44
Gp90	Hypothetical protein MycrhDRAFT_5728		53
Gp89	Hypothetical protein MycrhDRAFT_6888	Minor tail subunit	49
Gp88	Hypothetical protein MycrhDRAFT_6891	Lysin A	41
Gp85	Terminase	Terminase	72
Gp84	Hypothetical protein MycrhDRAFT_6894	Portal protein	82
	Hypothetical protein MycrhDRAFT_1970		31
	Hypothetical protein MycrhDRAFT_4121		29
	Hypothetical protein MycrhDRAFT_1415		
Gp83	Hypothetical protein MycrhDRAFT_6895		67
Gp82	Hypothetical protein MycrhDRAFT_6896	scaffold protein	59
Gp81	Hypothetical protein MycrhDRAFT_6897	Major head subunit	77
Gp80	Hypothetical protein MycrhDRAFT_6898		54
Gp79	Phage protein Gp19/Gp15/Gp42		74
Gp78	Hypothetical protein MycrhDRAFT_6901		77
Gp77	Hypothetical protein MycrhDRAFT_6902		66
Gp76	Hypothetical protein MycrhDRAFT_6903		70
Gp75	Hypothetical protein MycrhDRAFT_6904	Major tail subunit	73
Gp74	Hypothetical protein MycrhDRAFT_6905		50
Gp72	Hypothetical protein MycrhDRAFT_6906	Minor tail subunit (TMP)	32
Gp71	Hypothetical protein MycrhDRAFT_6907	Minor tail subunit	73
Gp70	Putative phage tail	Minor tail subunit	78
Gp69	hypothetical protein MycrhDRAFT_6909		73
Gp68	Bacteriophage protein		32
Gp64	Putative phage excisionase		72
	Hypothetical protein MycrhDRAFT_6242		30
	Hypothetical protein MycrhDRAFT_5633		32
Gp59	Hypothetical protein MycrhDRAFT_5784		58
Gp58	Hypothetical protein MycrhDRAFT_5783		47
Gp56	Hypothetical protein MycrhDRAFT_5782		42
Gp54	DNA-directed DNA polymerase	DNA polymerase	75
Gp53	Hypothetical protein MycrhDRAFT_5777	Transcriptional regulator LacI family	44
Gp52	Hypothetical protein MycrhDRAFT_5776		49
Gp51	Hypothetical protein MycrhDRAFT_5775	Probable DNA binding protein	53
Gp50	Hypothetical protein MycrhDRAFT_5774		52
Gp49	Thymidylate synthase thyX	Flavin-dependentthymidylate synthase (ThyX)	49
Gp48	Hypothetical protein MycrhDRAFT_5772	A class II ribonucleotidereductase (RNR)	55

The protein in SWU1	The protein in <i>M. rhodesiae</i> JS60	The predicted function of protein in SWU1	Identity
Gp47	Ribonucleoside-triphosphate reductase, adenosylcobalamin-dependent		79
Gp45	Hypothetical protein MycrhDRAFT_5770		77
Gp44	Hypothetical protein MycrhDRAFT_5769		51
Gp42	Metallophosphoesterase		68
Gp41	Hypothetical protein MycrhDRAFT_5767		48
Gp40	Glutaredoxin		60
	Glutaredoxin-like protein NrdH		36
Gp38	TOPRIM domain-containing protein		80
Gp35	Recombination endonuclease VII	Endonuclease	75
Gp33	Hypothetical protein MycrhDRAFT_5761		40
Gp30	Hypothetical protein MycrhDRAFT_5759		40
Gp29	Hypothetical protein MycrhDRAFT_5758		52
Gp28	Hypothetical protein MycrhDRAFT_5755	DnaB helicase	76
Gp25	Hypothetical protein MycrhDRAFT_5753		51
Gp24	Hypothetical protein MycrhDRAFT_5751		32
Gp23	Hypothetical protein MycrhDRAFT_5750		71
Gp22	Hypothetical protein MycrhDRAFT_5749		35
	Hypothetical protein MycrhDRAFT_4647		30
Gp20	Hypothetical protein MycrhDRAFT_5748	Repressor-like immunity protein(?)	57
Gp17	Hypothetical protein MycrhDRAFT_5730		41
Gp16	Hypothetical protein MycrhDRAFT_5742		63
Gp14	Hypothetical protein MycrhDRAFT_2544		26
Gp3	Hypothetical protein MycrhDRAFT_5733		65
Gp1	Hypothetical protein MycrhDRAFT_5729		45