Romulus and Remus, Two Phage Isolates Representing a Distinct Clade within the Twortlikevirus Genus, Display Suitable Properties for Phage Therapy Applications

Katrien Vandersteegen, Andrew M. Kropinski, John H. E. Nash, Jean-Paul Noben, Katleen Hermans, Rob Lavigne

Laboratory of Gene Technology, Division of Gene Technology, University of Leuven, Heverlee, Belgium; Public Health Agency of Canada, Laboratory for Foodborne Zoonoses, Guelph, Ontario, Canada; Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada; Biomedical Research Institute and Transnational University Limburg, School of Life Sciences, Hasselt University, Diepenbeek, Belgium; Department of Pathology, Bacteriology and Poultry Diseases, Ghent University, Merebeke, Belgium

The renewed interest in controlling Staphylococcus aureus infections using their natural enemies, bacteriophages, has led to the isolation of a limited number of virulent phages so far. These phages are all members of the Twortlikevirus, displaying little variance. We present two novel closely related (95.9% DNA homology) lytic myoviruses, Romulus and Remus, with double-stranded DNA (dsDNA) genomes of 131,333 bp and 134,643 bp, respectively. Despite their relatedness to Staphylococcus phages K, G1, ISP, and Twort and Listeria phages A511 and P100, Romulus and Remus can be proposed as isolates of a new species within the Twortlikevirus genus. A distinguishing feature for these phage genomes is the unique distribution of group I introns compared to that in other staphylococcal myoviruses. In addition, a hedgehog/intein domain was found within their DNA polymerase genes, and an insertion sequence-encoded transposase exhibits splicing behavior and produces a functional portal protein. From a phage therapy application perspective, Romulus and Remus infected approximately 70% of the tested S. aureus isolates and displayed promising lytic activity against these isolates. Furthermore, both phages showed a rapid initial adsorption and demonstrated biofilm-degrading capacity in a proof-of-concept experiment.

In addition to asymptomatic colonization of the skin and mucosal surfaces, Staphylococcus aureus is responsible for various infections of the skin, systemic infections, and sepsis (1–3). Like many other bacteria, S. aureus produces biofilms which are difficult to eradicate (4). In combatting this versatile opportunistic pathogen, bacteriophages have become an interesting alternative to antibiotics. Staphylococcus phages belong to the viral families Myoviridae, Podoviridae, and Siphoviridae, with the last being temperate so far and therefore not suitable for phage therapy.

The myoviruses infecting S. aureus appear to be genotypically and proteomically related and have been classified into the subfamily Spounavirinae, genus Twortlikevirus (5, 6). This genus includes staphylococcal phages K (7), G1, Twort (8), ISP, and A5-W display an intron distribution similar to those of K and G1 (13), whereas phage GH15 lacks all previously described introns (15). Although splicing in phages was first believed to be a regulatory mechanism of DNA metabolism (22, 23), the discovery of introns in late genes (8, 21, 24) challenges the relevance of this hypothesis. So far, no other solid prediction of the meaning of splicing in phages can be made. Moreover, the appearance of split genes is more plentiful in phages infecting Gram-positive bacteria (25).

In this report, we present a microbiological and molecular examination of two newly isolated phages, Romulus and Remus, as well as a proof-of-concept experiment which illustrates their biofilm-degrading potential. Furthermore, both phages carry introns, a hedgehog/intein domain, and an insertion sequence in their genome interrupting multiple open reading frames. Notably, Romulus and Remus represent a new species within the genus Twortlikevirus, which calls for expanding the existing taxonomy within this genus.
**MATERIALS AND METHODS**

**Bacterial isolates and culture conditions.** For phage enrichment purposes, a collection of *S. aureus* isolates was obtained from J. Verhaegen (Laboratory of Clinical Bacteriology and Mycology, Campus Gasthuisberg, UZ Leuven, Belgium) and the Institut Pasteur of Brussels (Belgium). Additional *Staphylococcus* isolates were provided by M. Vaneechoutte (Department of Clinical Chemistry, Microbiology and Immunology, Ghent University, Belgium), T. Ito (Department of Bacteriology, Juntendo University, Tokyo, Japan), the Centre for Food and Microbial Technology (University of Leuven, Belgium), the Department of Pathology, Bacteriology and Poultry Diseases (Ghent University, Belgium), A. Toledo-Arna (Instituto de Agrobiotecnologia and Departamento de Produccion Agraria, Universidad Publica de Navarra, Spain), and C. Loc-Carrillo (VA SLC Health Care System, Department of Orthopaedics, Salt Lake City, UT) for host range screening. Isolates were obtained in liquid Mueller-Hinton (MH) medium at 37°C, on solid MH medium (1.5% agar), or in MH soft agar overlays (0.35% agar).

**Sampling and bacteriophage enrichment.** Raw inlet sewage water was collected in sterile Falcon tubes and homogenized by shaking at 4°C, after which bacteria and large waste particles were removed by centrifugation (30 min at 3,000 × g and 4°C, Sorvall Legend RT +; Thermo Fisher Scientific Inc., Waltham, MA) and filtration (Durapore polycylinide difluoride [PVDF] membrane, 0.45 µm; Millipore Corporation, Billerica, MA). After enrichment (26) in liquid MH medium (overnight at 37°C) against a selected *S. aureus* host strain, chloroform was added and the bacterial debris was removed by centrifugation (30 min at 3,000 × g and 4°C). Finally, potential phages were collected by centrifugation (90 min at 28,000 × g and 4°C), and the resulting pellet was dissolved in phage buffer (10 mM Tris-Cl; 10 mM MgSO₄; 150 mM NaCl; pH 7.5) and spotted on a soft agar lawn containing the selected host. The presence of phages was evaluated after incubation (overnight at 37°C). In case of lysis spots, multiple single-plaque isolations were performed to obtain a pure phage isolate.

**Bacteriophage amplification and purification.** High-titer phage stocks were obtained through amplification in liquid MH medium. Following visible lysis of the liquid culture, the lysisate was incubated with chloroform for 10 min to kill any residual bacteria. Phages were collected through centrifugation (30 min at 3,000 × g and 4°C) Phage particles were purified and concentrated by polyethylene glycol (PEG) 8000 precipitation (27) and separated from contaminants by sedimentation in a CsCl (MP Biomedicals, Solon, OH) gradient (1.33, 1.45, 1.50, and 1.70 g/cm³ in SM buffer [50 mM Tris-Cl, 100 mM NaCl, 8 mM MgSO₄, 0.01% gelatin [Sigma-Aldrich]; pH 7.5]) in Ultra-Clear tubes (Beckman Coulter, Inc., Fullerton, CA). Finally, the phage preparation was dialyzed against phage buffer using Slide-A-Lyzer G2 dialysis cassettes (2K MWCO; Pierce, Rockford, IL) and stored at 4°C.

**Electron microscopic imaging.** Phage particles were collected by centrifugation at 25,000 × g for 1 h and washed twice in 0.1 M ammonium acetate (pH 7.0) using a Beckman high-speed centrifuge and a JA-18.1 fixed-angle rotor. Following deposition on carbon-coated copper grids and staining with 2% (wt/vol) potassium phosphotungstate (pH 7.0), phages were visualized in a Philips EM 300 transmission electron microscope (28).

**Host range screening.** The host range was analyzed by spotting serial dilutions of phage on a soft agar lawn of *S. aureus, Staphylococcus epidermidis*, and *Staphylococcus haemolyticus* isolates. The results were confirmed with a plaque assay which permitted assessment of the efficiency of plating, i.e., the relative phage titer on a bacterial strain compared to the maximum titer observed.

**Killing curve and adsorption curve.** For adsorption experiments, 100-µl phage samples were taken at fixed intervals (20 s and 1, 2, 5, 7, 10, 15, 20, and 25 min) and transferred to 850 µl of liquid MH medium supplemented with 50 µl of chloroform to lyse the remaining bacteria (26, 29). The number of unadsorbed or reversibly adsorbed phages was determined, and for each time point during the adsorption experiment, the adsorption rate constant was calculated. This value, k, is calculated by the equation \( \frac{2.3}{Bt} \left( \log \frac{P_0}{P} \right) \) where \( P \) represents the number of unadsorbed or reversibly adsorbed phages, \( P_0 \) the initial number of phages, B the bacterial titer, and t the time after infection.

To generate a killing curve at different multiplicities of infection (MOI; the ratio of phage particles to host cells), a bacterial culture was infected at an optical density at 600 nm (OD₆₀₀) of 0.3. The OD₆₀₀ of the infected cultures was monitored every 20 min for 6 h and compared with that of an uninfected bacterial culture.

**Biofilm degradation.** Biofilm degradation was examined using the MBEC assay (30, 31) in which 96 polystyrene pegs (Nunc-ImmuNo TSP; Thermo Fisher Scientific, Tournai, Belgium) are used to grow 96 identical biofilms. Overnight cultures grown in liquid MH medium were diluted (1:200) in 200 µl of MH medium supplemented with 1% (wt/vol) glucose. The peg lid was placed in the microtiter plate, sealed with Parafilm, and incubated at 37°C without shaking. Phage treatment of the biofilms occurred by placing the peg lid in 200 µl of phage buffer containing a specific amount of phages. To quantify the attached biofilms, the pegs were first washed briefly with phage buffer. Next, the biofilm mass was stained with 200 µl of a 0.1% (wt/vol) crystal violet (Merck) solution in 5% (vol/vol) isopropyl alcohol, 5% (vol/vol) methanol, and 90% (vol/vol) phosphate-buffered saline (PBS; 137 mM NaCl, 10 mM phosphate, 2.7 mM KCl; pH 7.4) for 30 min, washed again, and air dried for 30 min. The bound crystal violet was removed with 200 µl of 33% glacial acetic acid. The absorbance of crystal violet at 600 nm was measured using a Multiskan RC (Thermo Labsystems, Vantaa, Finland).

**Results of treatment with different phages and titers were compared to treatment with phage buffer using a two-tailed Student t test (P < 0.05).**

**Biophysical stability.** The pH stability of the phage particles was tested by incubation in pH buffer (150 mM potassium chloride, 10 mM potassium dihydrogen phosphate, 10 mM sodium citrate, 10 mM boric acid; adjusted to pH 1 to pH 13) for 24 h at room temperature. Titration of the phage samples and comparison with a control sample incubated in phage buffer permitted the calculation of the relative survival of the phages. Similarly, the temperature stability was examined by incubation in phage buffer at different temperatures (~20°C, 4°C, 16°C, 37°C, 42°C, and 50°C), with 4°C as a control.

**DNA isolation and genome sequencing.** Isolation of phage DNA was performed as described earlier (27). Sequencing of phage Remus was performed by Sanger shotgun sequencing and 454 sequencing (Plate-forme d’Analyses Génomiques at Laval University, Quebec, Quebec, Canada). Raw sequences arising from Sanger sequencing were filtered for vector contamination, and poor-quality reads were eliminated using the Pre判program (32). These were combined with the raw 454 sequence reads and assembled using the MIRA sequence assembly package, version 3.2.1.15 (33). The resulting assembly was imported into the Gap5 program (32), in which it was edited and proofread. Additional shotgun sequencing was done to elucidate ambiguities in A and T stretches. The phage Romulus genome was sequenced by 454 sequencing and upon comparison with that of phage Remus was verified with shotgun sequencing.

**In silico analysis.** Phage genomes were autoannotated using MyRAST (http://blog.theseed.org/servers/presentations/t1/running-a-job-with-the -desktop-rast.html) and then proofread in Kodon (Applied Maths, Austin, TX). Shine–Dalgarno sequences of these predicted open reading frames and potential alternative start codons were checked manually. Using BLASTP (34) and HHpred (35) with databases available in August 2011 and in June 2012, putative protein functions were assigned for Remus and Romulus, respectively. Promoters were predicted using MEME (36), followed by manual verification. TransTerm (37) and ARNold (38, 39) were used to detect potential terminators, and the free energy of their secondary structures was calculated using Mfold (40). tRNAs were predicted using tRNAscan-SE (41) and ARAGORN (42). A comparative genome figure was generated using CGView (43), and DNA homology between phages was examined with EMBoss stretcher (44, 45).
Proteome analysis. Phage proteins were dried by extraction with methanol-chloroform (1:0.75, vol/vol). The protein pellet was resuspended in loading buffer (1% [wt/vol] sodium dodecyl sulfate [SDS], 6% [wt/vol] sucrose, 100 mM 1,4-dithiothreitol, 10 mM Tris-HCl [pH 6.8], 0.0625% [wt/vol] bromophenol blue) (46) and loaded onto a 12% polyacrylamide gel. Following electrophoresis (SDS-PAGE) (1 h at 200 V), the gel was stained with Simply Blue SafeStain (Invitrogen) and fragmented and trypsinized according to the protocol of Shevchenko et al. (47). The eluted peptides were subjected to electrospray ionization tandem mass spectrometry (ESI-MS-MS) on a LCQ Classic (ThermoFinnigan, San Jose, CA) equipped with a nano-LC column switching system as described earlier (48). The resulting MS-MS data were analyzed using Mascot, version 2.3.01, and Sequest, version 1.2.0.208, against a local database of all possible phage proteins.

Nucleotide sequence accession numbers. The nucleotide sequences of Romulus and Remus have been deposited in the NCBI database under accession number JX846613.

RESULTS

Isolation and morphology. Three phage propagation strains (Staphylococcus aureus PS47, PS92, and PSPB) and two clinical isolates (S. aureus KS8 and KS13) were selected for phage enrichment. Both clinical isolates carry genes conferring resistance to methicillin, penicillin, tetracycline, and aminoglycosides (11). Analysis of 23 sewage samples of the UZ Leuven Campus Gasthuisberg (Belgium), collected in January 2009, resulted in successful phage enrichment against PS47 and PSPB. In each case, plaques were about 1 mm in diameter and clear. Restriction analysis of both phage isolates (data not shown) revealed two similar but slightly different patterns, implying the isolation of two closely related lytic phages. The phages were named after Romulus and Remus, twin brothers and founders of Rome.

Based on their morphology examined by transmission electron microscopy (Fig. 1), Romulus and Remus were classified as members of the family Myoviridae. Both phages possess an isometric head with a diameter of 90 nm and a contractile tail with a length of 204 nm and a width of 17 nm. These dimensions correspond to those of other twortlikeviruses infecting S. aureus (5). Consequently, phage SA11, closely related to Romulus and Remus, was presumably erroneously described as a siphovirus (14).

Host range screening. A collection of 90 S. aureus isolates, 9 S. haemolyticus isolates, and 1 S. epidermidis isolate was used to assess the host range and corresponding efficiency of plating (see Table S1 in the supplemental material). For completeness, the host range of phage ISP was added as additional isolates were examined compared to earlier results (11). Romulus and Remus were able to infect 69% and 68% of all tested S. aureus isolates, respectively, which is considerably lower than the 87% infected by ISP. Further, while all human S. aureus isolates were sensitive to ISP infection, Romulus and Remus infected 28 and 30 out of 36 S. aureus isolates from patients, respectively. In contrast to the 21 phage propagation strains infected by ISP, only 19 and 18 of the 31 phage propagation strains were sensitive to Romulus and Remus, respectively. However, all the pig isolates which were insensitive to ISP infection were killed by Romulus and Remus. Similar to the case with ISP (11), none of the S. haemolyticus or S. epidermidis isolates showed sensitivity to Romulus and Remus infection. In general, Romulus and Remus displayed a moderate to low efficiency of plating, while a notable number of human isolates was infected with a high efficiency of plating by Remus. S. aureus isolates from rabbits are infected with a high efficiency of plating by all three phages.

Infection parameters. To examine the infection parameters, the adsorption to and subsequent infection of Staphylococcus aureus subsp. aureus Rosenbach ATCC 6538 at different MOIs in liquid culture were studied. Adsorption experiments were performed in the presence of approximately 3 × 10^8 bacterial cells/ml. The adsorption assay (data not shown) indicated that approximately 60% of the Romulus and Remus particles were irreversibly adsorbed within 1 min, with adsorption rate constants of 2.1 × 10^-9 ml/min and 2.2 × 10^-9 ml/min, respectively. Furthermore, for both phages 80% of the viral particles were bound after 25 min, with adsorption rate constants of 2.6 × 10^{-11} ml/min and 3.7 × 10^{-11} ml/min, respectively. These adsorption rate constants illustrate an initial rapid adsorption followed by a reduced rate leading to maximum adsorption, as was already observed for Escherichia phage T2 in 1961 (49).

Romulus and Remus showed a similar infection pattern on S. aureus subsp. aureus Rosenbach ATCC 6538 (data not shown) and were both able to kill the culture at MOIs of 1 and 10. In case of both phages, new phage particles were liberated after a lytic cycle of approximately 40 min. At an MOI of 0.1 the optical density at 600 nm of the bacterial culture decreased, but not to a steady minimum within the time period of the experiment. In contrast, S. aureus subsp. aureus Rosenbach ATCC 6538 was only attenuated by ISP (11). Considering therapeutic implementation, killing curves were generated as well for S. aureus KS23 isolated from a patient at the UZ Leuven Campus Gasthuisberg (11). Again, the optical densities upon infection with Romulus and Remus progressed in parallel (data not shown). Further, the bacterial culture

![FIG 1 Morphology of Romulus and Remus. Shown are transmission electron microscopic images of myoviruses Romulus (A) and Remus (B) negatively stained with 2% (wt/vol) potassium phosphotungstate (pH 7.0). Scale bars represent 100 nm.](http://jvi.asm.org/)
was killed for all three MOIs tested, while ISP could only cause a decrease in the optical density of which the minimum was not reached within the time period of the experiment (data not shown).

Biofilm degradation: proof-of-concept experiment. Before a proof-of-concept experiment was carried out to illustrate the biofilm-degrading potential of phages ISP, Romulus, and Remus, S. aureus PS47—a phage propagation strain characterized by a high relative biofilm-forming capacity (data not shown) and sensitivity to all three phages (see Table S1 in the supplemental material)—was selected. Using the MBEC assay (30, 31), we evaluated the effects of ISP, Romulus, and Remus on a 24-h pregrown biofilm of S. aureus PS47 (Fig. 2). Phage treatment for 2 to 6 h did not result in any substantial biofilm degradation, in contrast to when phages were applied for 24 h. The effects of ISP and Romulus treatments were significant only when 10^9 phages per peg were applied, while 10^7 to 10^9 Remus particles per peg resulted in a significant decrease of biofilm mass. When applied at 10^7 phages per peg, ISP, Romulus, and Remus were able to degrade the S. aureus PS47 biofilm by approximately 37.8%, 34.4%, and 60.4%, respectively. S. aureus Xen29, used to examine the biofilm-degrading activity of phage K (19), is considered to display an intermediate biofilm-degrading capacity using the MBEC assay in this study (data not shown).

Biophysical stability. To determine optimal storage and application conditions, the survival of Romulus and Remus at different temperatures and pH values was tested. Both phages appeared to be stable at 4°C and 16°C. At 37°C, the phage titers dropped by approximately 1 logarithmic unit, and at 42°C and 50°C, the logarithmic decrease was considerable. This should be taken into account in case of medical use. The pH range 5 to 9 caused no loss of phage infectivity, while a pH of 10 led to a logarithmic reduction of 1.5 and 4.5 units for Romulus and Remus, respectively. Like phage ISP (11), Romulus and Remus are therefore not suitable for oral administration without manipulation.

Genome analysis. (i) General features. The linear double-stranded DNA of Remus with terminally redundant ends comprises 134,643 bp (JX846612) encoding 189 putative ORFs (Fig. 3; see also Table S2 in the supplemental material). The 131,333-bp genome of Romulus (JX846613), on the other hand, is highly similar to that of Remus, except for a 4.4-kb deletion resulting in the absence of 12 putative ORFs (ORF123 to ORF134), a 1.1-kb insertion leading to three additional putative genes (ORF140A to ORF140C), and two point mutations (see Table S3 in the supplemental material). Forty-three of these predicted genes are transcribed from the minus strand, including the lysis module. A combination of prediction programs followed by manual verification predicted 61 and 60 promoters and 26 and 25 terminators for Remus and Romulus, respectively. Furthermore, one tRNA encoding tRNA^{Ser} was detected in an intergenic region. In addition, Romulus and Remus do not encode virulence-associated or toxic proteins and are therefore potentially safe to use in phage therapy.

(ii) Comparative genome analysis. Bioinformatic analysis of the Romulus and Remus genomes revealed 95.9% identity at the nucleotide level. A DNA comparison of both phage genomes with the genomes of K (AY176327), G1 (AY954969), Twort (NC_007021), Sb-1 (HQ163896), and ISP (FR852584) and the recently published phages Staph1N (IX080300), Fi200W (IX080303), P4W (IX080305), 676Z (IX080302), A3R (IX080304), MSA6 (IX080304), A5-W (EU418428), SA11 (JX194239), and GH15 (JQ686190) is represented in Table 1. The homology with the closely related phages (>90% nucleotide homology) K, G1, Sb-1, and ISP is as low as the homology between K, G1, Sb-1, and ISP on one hand and Twort on the other hand (approximately 55%). Furthermore, SA11 and the eight phages described by Łobocka et al. (13) and Gu and coworkers (15) (nucleotide identity between 88.6% and 99.9%) display a nucleotide identity between 53.2% and 58.7% to Romulus and Remus. In addition,
Romulus, Remus, and SA11 share only 53.2%, 53.7%, and 55.8% DNA identity, respectively, with phage Twort. Figure 3 shows a BLASTN comparison of the genomes of phages Remus, K, and Twort. Despite this relatively low DNA similarity, Romulus and Remus display the same modular genome organization as K, G1, Twort, and ISP. Three major modules encoding virion proteins, DNA metabolism functions, and proteins necessary for cell lysis can be distinguished. Predicted structural proteins of Romulus, Remus, and SA11 show approximately 70 to 80% sequence identity to proteins of the other twortlikeviruses infecting *S. aureus* and approximately 60% identity to proteins of the *Listeria* phages A511 and P100 (5, 50, 51). In general, other proteins encoded by the Romulus and Remus genomes displayed an average 50% homology with proteins of K, G1, ISP, and Twort. For 60 and 58% of the proteins predicted for Remus and Romulus, no function could be assigned using BLASTP or HHpred, and for 109 and 118 proteins, no database counterpart was found at all. The lysis module, for example, appeared to be significantly different from the module observed in phages K, G1, ISP, and Twort. While the holins of Romulus and Remus are homologous to the Twort holin, the endolysin, interrupted by a group I intron in K, G1, and ISP, is replaced by two presumed endolysins homologous to siphoviral endolysins in Romulus and Remus.

(iii) Genes interrupted by self-splicing elements. The presence of mobile elements in essential genes of Romulus and Remus is a characteristic feature among most known staphylococcal myoviruses (7,8, 11, 13). In the case of Romulus and Remus, six protein-encoding genes were found to be interrupted by six coding and two noncoding introns, an insertion sequence, and a hedgehog/intein domain. The distribution of the coding introns, however, is dissimilar from those observed before in *S. aureus* phages (7, 8, 11, 13).

First, the terminase large-subunit gene is fragmented by three group I introns of which the first and the third encode a homing endonuclease (Fig. 4A). The three introns were shown to be unrelated, whereas the embedded homing endonucleases, named I-RoReI and I-RoReII, appeared to be mutually homologous, implying a common evolutionary ancestor. In addition, HHpred analysis revealed relatedness of these homing endonucleases with...
An insertion sequence comprising a transposase gene (ORF14) and an additional gene (ORF13) was detected between ORF12 and ORF15, both homologous to parts of the gene encoding the portal protein in K, G1, ISP, and Twort (Fig. 4B). The transposase gene is preceded by the predicted promoter P5. According to BLAST analysis, the transposase in Romulus and Remus is homologous to the IS605 family transposase OrfA. The insertion sequences of this family typically comprise two genes encoding OrfA and OrfB, serving together as the functional transposase (53). ORF13 and ORF14 are transcribed from the same strand, whereas the IS605 family gene encoding OrfA is transcribed from the reverse strand (54). This raises the question of whether ORF13 is part of the insertion sequence. This question is strengthened by the imperfect palindromic sequences (29 to 30 nucleotides [nt]), which replace the inverted terminal repeats in the IS605 family of insertion sequences (55), and the direct target repeats (7 nt) flanking ORF14.

Third, the helicase gene of Romulus and Remus is split into two open reading frames, ORF49 and ORF50 (Fig. 4C). The insert is approximately 245 nucleotides in length and probably constitutes a group I intron lacking an open reading frame. The interruption occurs at the same location as where a hedgehog/intein (Hint) domain is found in the helicase gene of phage T4 (8, 13).

Next, the ribonucleotide (or ribonucleoside disphosphate) reductase large-subunit gene is interrupted by two group I introns (Fig. 4D). Again, the introns are not related. The embedded homing endonucleases, I-RoRelIII and I-RoRelIV, are homologous to I-TevI and I-HmuI, members of the GIY-YIG and HNH families, respectively (56, 57) (Table 2). The ribonucleotide reductase large-subunit gene itself is related to the Polaribacter r Igoriensii 23-P ribonucleotide reductase large-subunit gene and does not show any protein homology with the ribonucleotide reductase large-subunit gene of phages K, G1, Twort, ISP, Sb-1, A5-W, Staph1N, Fi200W, P4W, 676Z, A3R, or MSA6.

### TABLE 1 Comparative nucleotide analysis between genomes of *S. aureus* phages present in the NCBI database

<table>
<thead>
<tr>
<th>Phage</th>
<th>K</th>
<th>G1</th>
<th>Twort</th>
<th>ISP</th>
<th>A5-W</th>
<th>Fi200W</th>
<th>P4W</th>
<th>676Z</th>
<th>A3R</th>
<th>MSA6</th>
<th>Sb-1</th>
<th>GH15</th>
<th>SA11</th>
<th>Romulus</th>
<th>Remus</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Nucleotide identity</td>
<td>100</td>
<td>90.3</td>
<td>56.1</td>
<td>90.6</td>
<td>96.1</td>
<td>89.2</td>
<td>89.4</td>
<td>93.5</td>
<td>88.6</td>
<td>96.6</td>
<td>80.0</td>
<td>58.4</td>
<td>57.6</td>
<td>57.9</td>
<td></td>
</tr>
</tbody>
</table>

*The nucleotide identity between the genomes of each set of two phages was calculated using EMBOSS stretcher. Homologies lower than 60% are shaded. Complete (100%) homology is indicated by bold type.*

### TABLE 2 Overview of the predicted homing endonucleases embedded in group I introns in the Romulus and Remus genomes

<table>
<thead>
<tr>
<th>Homing endonuclease</th>
<th>ORF</th>
<th>Target gene</th>
<th>Homolog</th>
<th>Homing endonuclease family</th>
<th>Prediction program</th>
<th>E value</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-RoRel</td>
<td>ORF2</td>
<td>Terminus large subunit</td>
<td>I-Bth0305I (B. thuringiensis phage 0305φ-36)</td>
<td>EDXHD</td>
<td>HHpred</td>
<td>6.00E−17</td>
</tr>
<tr>
<td>I-RoRelII</td>
<td>ORF5</td>
<td>Terminus large subunit</td>
<td>I-Bth0305I (B. thuringiensis phage 0305φ-36)</td>
<td>EDXHD</td>
<td>HHpred</td>
<td>6.00E−17</td>
</tr>
<tr>
<td>I-RoRelIII</td>
<td>ORF67</td>
<td>Ribonucleotide reductase large subunit</td>
<td>I-TevI (E. coli phage T4)</td>
<td>GIY-YIG</td>
<td>BLASTP/HHpred</td>
<td>6.00E−10/1.10E−24</td>
</tr>
<tr>
<td>I-RoRelIV</td>
<td>ORF69</td>
<td>Ribonucleotide reductase large subunit</td>
<td>I-HmuI (B. subtilis phage SPO1)</td>
<td>HNH</td>
<td>HHpred</td>
<td>7.20E−40</td>
</tr>
<tr>
<td>I-RoReV</td>
<td>ORF76</td>
<td>DNA polymerase</td>
<td>I-HmuI (B. subtilis phage SPO1)</td>
<td>HNH</td>
<td>HHpred</td>
<td>1.70E−21</td>
</tr>
<tr>
<td>I-RoReVI</td>
<td>ORF81</td>
<td>DNA repair protein</td>
<td>I-Dml (Desulfurococcus mobilis)</td>
<td>LAGLIDADG</td>
<td>HHpred</td>
<td>2.50E−24</td>
</tr>
</tbody>
</table>

*For each predicted homing endonuclease, the encoding ORF and the target gene are given. Further, the homologous homing endonuclease, its family, the prediction program, and corresponding E value are indicated.*

---

Bacillus thuringiensis phage 0305-36 I-Bth0305I, the first and until now only member of the EDXHD family of homing endonucleases discovered in a phage genome (52) (Table 2). Furthermore, I-RoRel and I-RoRelI share homology with the homing endonucleases residing within the group I intron in the terminase large-subunit genes of *Staphylococcus* phages MSA6 (1-MSal) and Twort (8, 13).

An insertion sequence comprising a transposase gene (ORF14) and an additional gene (ORF13) was detected between ORF12 and ORF15, both homologous to parts of the gene encoding the portal protein in K, G1, ISP, and Twort (Fig. 4B). The transposase gene is preceded by the predicted promoter P5. According to BLAST analysis, the transposase in Romulus and Remus is homologous to the IS605 family transposase OrfA. The insertion sequences of this family typically comprise two genes encoding OrfA and OrfB, serving together as the functional transposase (53). ORF13 and ORF14 are transcribed from the same strand, whereas the IS605 family gene encoding OrfA is transcribed from the reverse strand (54). This raises the question of whether ORF13 is part of the insertion sequence. This question is strengthened by the imperfect palindromic sequences (29 to 30 nucleotides [nt]), which replace the inverted terminal repeats in the IS605 family of insertion sequences (55), and the direct target repeats (7 nt) flanking ORF14.

Third, the helicase gene of Romulus and Remus is split into two open reading frames, ORF49 and ORF50 (Fig. 4C). The insert is approximately 245 nucleotides in length and probably constitutes a group I intron lacking an open reading frame. The interruption occurs at the same location as where a hedgehog/intein (Hint) domain is found in the helicase gene of phage T4 (8, 13).

Next, the ribonucleotide (or ribonucleoside disphosphate) reductase large-subunit gene is interrupted by two group I introns (Fig. 4D). Again, the introns are not related. The embedded homing endonucleases, I-RoRelIII and I-RoRelIV, are homologous to I-TevI and I-HmuI, members of the GIY-YIG and HNH families, respectively (56, 57) (Table 2). The ribonucleotide reductase large-subunit gene itself is related to the Polaribacter r Igoriensii 23-P ribonucleotide reductase large-subunit gene and does not show any protein homology with the ribonucleotide reductase large-subunit gene of phages K, G1, Twort, ISP, Sb-1, A5-W, Staph1N, Fi200W, P4W, 676Z, A3R, or MSA6.
The DNA polymerase gene not only is interrupted by a group I intron but also contains a Hint domain, an autoprocessing domain which undergoes a protein splicing reaction (Fig. 4E). This Hint domain is located at the end of ORF75, with predicted N- and C-terminal boundaries at amino acid positions 732 and 1140, and typically contains a histidine and an asparagine residue at the C terminus, while a threonine residue follows the downstream splice site. The group I intron...
codes for a homing endonuclease, I-RoReVI, which is also related to I-HmuI (Table 2).

The sixth group 1 intron-invaded gene is the gene coding for a DNA repair protein (Fig. 4F). This intron codes for I-RoReVI, an endonuclease similar to I-DmoI, an archaeal intron-encoded homing endonuclease belonging to the LAGLIDADG family (Table 2); BLASTP analysis showed homology with other members of this family as well (data not shown). Remarkably, an LAGLIDADG homing endonuclease has never been reported for a phage genome before. The homolog of this DNA repair protein in Staphylococcus phages MSA6 and 812 is interrupted by a group I intron as well (13, 16). In addition, the embedded homing endonuclease in Romulus and Remus, I-RoReVI, resembles the corresponding homing endonuclease in MSA6 and 812 (13,16).

It should be mentioned that the above-mentioned group I introns were predicted with bioinformatic tools and therefore require experimental evidence.

**Structural proteome analysis of Remus.** Mass spectrometric analysis of 20 gel slices of the Remus structural proteome resulted in the identification of 19 proteins (Table 3). Homologs of 18 of these proteins were already identified in earlier proteome studies (11, 59), while Gp42 of Remus can now be considered a component of the virus particle as well. Various Remus proteins were detected throughout the SDS-PAGE gel, illustrated by the major capsid (Gp18) and the major tail sheath (Gp25) protein, suggesting their abundant presence in the Remus virion. Similar observations were made for phage ISP (11).

The five major protein bands (Fig. 5) of the Remus structural proteome are very similar to the bands visualized for phage ISP (11). However, band D of Remus is localized several kilodaltons lower. Because for ISP no corresponding molecular mass was found, this band was thought to be a proteolytically processed form of the tail sheath protein (59, 60). If this hypothesis is correct,

---

**TABLE 3 Characteristics of the ESI-MS-MS-identified proteins of bacteriophage Remus**

<table>
<thead>
<tr>
<th>Gp</th>
<th>Band no.</th>
<th>Molecular mass (kDa)</th>
<th>No. of identified peptides</th>
<th>Sequence coverage (%)</th>
<th>HHpred search</th>
<th>E value</th>
<th>ISP homolog</th>
<th>Molecular mass (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gp12</td>
<td>7</td>
<td>33.8</td>
<td>4</td>
<td>15.0</td>
<td>Portal protein (HK97)</td>
<td>5.8E−32</td>
<td>Gp8</td>
<td>64.1</td>
</tr>
<tr>
<td>Gp15</td>
<td>7</td>
<td>26.1</td>
<td>4</td>
<td>19.7</td>
<td>Portal protein (HK97)</td>
<td>2.2E−30</td>
<td>Gp8</td>
<td>64.1</td>
</tr>
<tr>
<td>Gp16</td>
<td>19</td>
<td>28.8</td>
<td>2</td>
<td>9.8</td>
<td>Prohead protease (HK97)</td>
<td>9.5E−34</td>
<td>Gp9</td>
<td>28.6</td>
</tr>
<tr>
<td>Gp18</td>
<td>11, 2, 4–12</td>
<td>51.0</td>
<td>16</td>
<td>52.9</td>
<td>Major capsid protein (HK97)</td>
<td>5.2E−05</td>
<td>Gp11</td>
<td>51.2</td>
</tr>
<tr>
<td>Gp21</td>
<td>11</td>
<td>33.6</td>
<td>2</td>
<td>9.3</td>
<td></td>
<td></td>
<td>Gp14</td>
<td>33.7</td>
</tr>
<tr>
<td>Gp25</td>
<td>5–9, 11, 14–16, 18–20</td>
<td>64.6</td>
<td>24</td>
<td>51.9</td>
<td>Tail sheath protein (Mu)</td>
<td>3.1E−61</td>
<td>Gp18</td>
<td>64.5</td>
</tr>
<tr>
<td>Gp26</td>
<td>18, 19</td>
<td>12.4</td>
<td>4</td>
<td>55.4</td>
<td>Tail tube protein (T4)</td>
<td>0.74</td>
<td>Gp19</td>
<td>15.9</td>
</tr>
<tr>
<td>Gp36</td>
<td>14</td>
<td>34.9</td>
<td>1</td>
<td>5.7</td>
<td>Cell wall peptidase (NlpC)</td>
<td>2.7E−24</td>
<td>Gp28</td>
<td>34.6</td>
</tr>
<tr>
<td>Gp38</td>
<td>13</td>
<td>30.9</td>
<td>9</td>
<td>36.0</td>
<td>Baseplate structural protein (T4)</td>
<td>96</td>
<td>Gp30</td>
<td>29.3</td>
</tr>
<tr>
<td>Gp39</td>
<td>16</td>
<td>20.1</td>
<td>2</td>
<td>13.3</td>
<td></td>
<td></td>
<td>Gp31</td>
<td>20.0</td>
</tr>
<tr>
<td>Gp41</td>
<td>9</td>
<td>39.2</td>
<td>3</td>
<td>9.2</td>
<td>Baseplate J-like protein (P2)</td>
<td>8.5E−43</td>
<td>Gp33</td>
<td>39.2</td>
</tr>
<tr>
<td>Gp42</td>
<td>3, 5</td>
<td>102.7</td>
<td>6</td>
<td>9.0</td>
<td>Tail protein (P2)</td>
<td>1.6E−4</td>
<td>Gp34</td>
<td>116.3</td>
</tr>
<tr>
<td>Gp43</td>
<td>17</td>
<td>19.2</td>
<td>3</td>
<td>20.8</td>
<td>Baseplate structural protein (T4)</td>
<td>3.7E−35</td>
<td>Gp35</td>
<td>19.2</td>
</tr>
<tr>
<td>Gp44</td>
<td>1, 2, 4–6</td>
<td>129.4</td>
<td>20</td>
<td>24.9</td>
<td></td>
<td></td>
<td>Gp36</td>
<td>129.1</td>
</tr>
<tr>
<td>Gp46</td>
<td>6, 14</td>
<td>72.8</td>
<td>5</td>
<td>9.7</td>
<td></td>
<td></td>
<td>Gp38</td>
<td>72.6</td>
</tr>
<tr>
<td>Gp58</td>
<td>15, 16, 18, 19</td>
<td>23.4</td>
<td>8</td>
<td>61.4</td>
<td></td>
<td></td>
<td>Gp70</td>
<td>23.2</td>
</tr>
<tr>
<td>Gp86</td>
<td>16, 18–20</td>
<td>18.2</td>
<td>6</td>
<td>54.7</td>
<td>Major tail protein (lambda)</td>
<td>9.5E−21</td>
<td>Gp71</td>
<td>17.9</td>
</tr>
<tr>
<td>Gp87</td>
<td>20</td>
<td>7.6</td>
<td>2</td>
<td>50.0</td>
<td></td>
<td></td>
<td>Gp72</td>
<td>7.8</td>
</tr>
<tr>
<td>Gp97</td>
<td>19</td>
<td>17.7</td>
<td>3</td>
<td>19.7</td>
<td></td>
<td></td>
<td>Gp83</td>
<td>17.8</td>
</tr>
</tbody>
</table>

For each gene product (Gp), the band number(s) (according to Fig. 5), the molecular mass, the number of identified peptides (>99% protein identification probability with manual validation), and the protein sequence coverage are listed. Furthermore, the putative protein function according to HHpred and the corresponding E value are shown, as well as homologous ISP proteins and their molecular masses.
the major tail sheath protein of Remus (Gp25), of which the molecular mass corresponds to that of the major tail sheath protein of ISP, is modified in a different way. For Gp12 as well as Gp15 of Remus, both corresponding to the portal proteins of K, G1, and ISP encoded by only one open reading frame, peptides were detected in band number 7 (Table 3).

A thorough HHpred analysis in which a higher E value was also taken into account (Table 3) permitted us to confirm and even specify the function of several identified structural proteins. For several Remus proteins, homology with *Escherichia* phages T4, HK97, and P2 was detected. Moreover, the tail protein Gp26 is homologous to the tail tube protein of T4, and Gp38, Gp41, and Gp43 are proposed to be components of the Remus baseplate. The molecular masses of Gp26 (12.4 kDa) and Gp42 (102.7 kDa) are markedly lower than the 15.9 kDa and 116.3 kDa of the putative tail tube protein and corresponding tail protein of phage ISP, probably resulting in different dimensions of the contractile myoviral tail.

**DISCUSSION**

The presence of the pathogenic bacterium *S. aureus* is often associated with hospital environments (61, 62), while bacteriophages are expected to be widespread in locations populated by their bacterial host. Therefore, this research focused on hospital sewage for phage isolation, resulting in the isolation of the novel phages Romulus and Remus. Phage isolation attempts in this research, however, have revealed that the isolation of new virulent *Staphylococcus* phages is not a straightforward approach, illustrated by the limited number of virulent *Staphylococcus* phages described in the scientific literature (7–9, 11–15, 63, 64).

Like the majority of the known virulent *Staphylococcus* phages, Romulus and Remus were assigned to the genus *Twortlikevirus* based on their morphological features, including a characteristic baseplate at the end of the tail (5, 13). Furthermore, both phages were shown to infect a considerable number of *S. aureus* isolates, which was also observed for *twortlikeviruses* K, ISP, Sb-1, and MSA6 (9, 11, 12, 17). From a genome perspective, Romulus and Remus are very closely related to each other and to phage SA11 but are substantially different from the other *twortlikeviruses* infecting *S. aureus*. At the DNA level, Romulus, Remus, and SA11 display no more than 60% homology with these staphylococcal myoviruses, suggesting that they represent a new species within *Twortlikevirus* (Fig. 3). This places Romulus, Remus, and SA11 in a new clade within *Twortlikevirus* (representative phage, Remus), next to and clearly divergent from phages K, G1, ISP, A5-W, Staph1N, Fi200W, P4W, 676Z, A3R, MSA6, and Sb-1 on one hand (representative phage, K) and phage Twort on the other hand (representative phage, Twort).

Although *Staphylococcus* phages MSA6, Twort, and 812 are quite distantly related to Romulus and Remus at the DNA level, these phages contain similar group I intron-embedded homing endonucleases. Interestingly, the related homing endonucleases target functional homologs. In addition to the ribonucleotide reductase large-subunit gene of *Bacillus subtilis* bacteriophage SPB (24), the DNA polymerase gene of Romulus and Remus provides a second example of a group I intron and an intein coding sequence within the same gene. Furthermore, an insertion sequence in Romulus and Remus is unusually located between ORF12 and ORF15, genes encoding parts of the portal protein. For Gp12 and Gp15, peptides were detected in the same band. This implies that both genes are transcribed and translated into one mature, functional protein, suitable to be incorporated in the phage particle despite the presence of an insertion sequence. For the first time, evidence was provided for intron-like splicing behavior of an insertion sequence-encoded transposase.

Despite their close relatedness, the differences in host range and efficiency of plating between Romulus and Remus are remarkable. Their host range is, however, relatively narrow compared to that of phage ISP, known for its broad host range, including human clinical isolates (10, 11). Nonetheless, the *in vitro* infection patterns of Romulus and Remus appeared more promising than those of ISP, implying their suitability for efficient control of sensitive *S. aureus* isolates. Furthermore, the biofilm-degrading potential of Romulus and Remus makes their therapeutic application even more attractive. In consideration of their therapeutic potential, the temperature and pH range in which Romulus and Remus are stable were determined.

To conclude, Romulus and Remus are considered to be members of a new species within *Twortlikevirus*, containing a unique scale of group I introns, an intein, and an insertion sequence in their genomes. In addition, the characterization of both phage isolates, comprising microbiological as well as molecular aspects, demonstrated their appropriateness for therapeutic application.

**ACKNOWLEDGMENTS**

This work was supported by the Federal Public Service of Health, Food Chain Safety and Environment, Belgium (contract RT 08/6 PHAGE). We thank H.-W. Ackermann for performing the electron microscopic imaging.

**REFERENCES**


