Genomic characterization of lytic *Staphylococcus aureus* phage GH15: providing new clues to intron shift in phages

Jingmin Gu, Xiaohe Liu, Mei Yang, Yue Li, Changjiang Sun, Rong Lu, Jun Song, Qingming Zhang, Liancheng Lei, Xin Feng, Chongtao Du, Hao Yu, Yongjun Yang and Wenyu Han

Phage GH15 is a polyvalent phage that shows activity against a wide range of *Staphylococcus aureus* strains. This study analysed the genome of GH15. The genome size of GH15 (139 806 bp) was found to be larger than that of the known staphylococcal phages, and the G+C content (30.23 mol%) of GH15 was lower than that of any other staphylococcal myovirus phages. By mass spectrometry, ten structural proteins were identified. Analysis revealed that GH15 was closely related to phages G1, ISP, A5W, Sb-1 and K, and was moderately related to Twort. In light of the variability in identity, coverage, G+C content and genome size, coupled with the large number of mosaicisms, there certainly were close evolutionary relationships from K to Sb-1, A5W, ISP, G1 and finally GH15. Interestingly, all the introns and inteins present in the above phages were absent in GH15 and there appeared to be intron loss in GH15 compared with the intron gain seen in other phages. A comparison of the intron- and intein-related genes demonstrated a clear distinction in the location of the insertion site between intron-containing and intron-free alleles, and this might lead to the establishment of a consensus sequence associated with the presence of an intron or intein. The comparative analysis of the GH15 genome sequence with other phages not only provides compelling evidence for the diversity of staphylococcal myovirus phages but also offers new clues to intron shift in phages.

**INTRODUCTION**

Bacteriophages are the most abundant and diverse life form. The lysogenic life cycle of a phage is thought to exert an important influence on both the evolution of both the phage and the host (Krupovic et al., 2011). Moreover, phages play critical roles not only in bacterial ecology and adaptation to new environments but also in the pathogenesis of human bacterial infections (Kwan et al., 2005). *Staphylococcus aureus* is an important pathogen that underlies various acute and chronic infectious diseases. The pathogenicity of *S. aureus* is related in some cases to phages and pathogenicity islands (Goerke et al., 2009; Tormo-Más et al., 2010). In particular, phages play an important role in the pathogenicity of *S. aureus* either by carrying accessory virulence factors or by interrupting chromosomal virulence genes (Goerke et al., 2009).

In recent years, a number of staphylococcal phages have been isolated and studied. Kwan et al. (2005) reported that staphylococcal phages can be organized into three classes based on genome size, the organization of the gene map and comparative nucleotide and protein sequence analyses. Class III consists of phages with large genomes (>125 kb) (K, A5W, ISP, Sb-1, G1 and Twort). All phages of this class can be structurally defined as members of the family *Myoviridae*. In fact, class III is consistent with the genus of ‘Twort-like viruses’ based on proteomic data and physicochemical parameters (Lavigne et al., 2009).

Previous studies have illustrated that there are introns and/or inteins in class III phages (Kwan et al., 2005). The predicted lysin and polymerase genes of phages from this class, excluding Twort, are interrupted by one or more introns (O’Flaherty et al., 2004). In addition, several other genes of Twort have been reported to contain introns or inteins (Kwan et al., 2005; Landthaler et al., 2002; Landthaler & Shub, 1999). Recently, we isolated a novel staphylococcal myovirus phage, named GH15. The lytic ability of the GH15 lysin protein and the binding activity of its SH3b domain have been described previously (Gu et al., 2011a, b, c). Interestingly, GH15 lysin shares a high degree of similarity with the lysins of most phages belonging to class III, but there is no intron in the lysin gene of GH15. Thus, although GH15 appears to be closely related to other members of the staphylococcal class III phages, it possesses...
clear differences. The GH15 genome has recently been sequenced (Gu et al., 2012), allowing further analysis of its genome structure.

In this paper, we have reported the results of genomic analysis of this new staphylococcal phage. Elucidation of the structure of this phage genome may provide insights into the diversity of staphylococcal myobacteriophages and the process of intron shift in these phages.

RESULTS AND DISCUSSION

General features of the GH15 genome

As illustrated in Fig. S1 (available in JGV Online), the contractibility of the tail of GH15 indicated that morphologically GH15 is a member of the family Myoviridae, similar to phages K and 812 (Pantácek et al., 1998; Rees & Fry, 1981). The complete genome of GH15 was analysed with a variety of bioinformatics tools, and was found to be a 139 806 bp contiguous sequence of linear dsDNA. The complete genome of GH15 encoded 214 predicted ORFs. The gene-coding potential of the global genome was 89.37 %, with a mean ORF size of 653 bp. The analysis indicated that 129 predicted proteins (60.28 %) with >100 residues were encoded by the GH15 genome, and their arrangement at the whole-genome level was mapped (Fig. 1). Of all ORFs, 196 (91.6 %) initiated translation with an AUG start codon, whereas only ten (4.7 %) initiated translation with a UUG codon and eight (3.7 %) with a GUG codon (Table S1).

The whole genome could be divided into two distinct regions, which were transcribed divergently. The majority of genes (150 of 214 ORFs) were transcribed in one orientation, and the rest were transcribed in the opposite orientation, with most of the latter grouped together and located in the region from 6.5 to 41 kb (ORF17–ORF78). All predicted proteins were examined for similarity to known bacterial and phage sequences deposited in public databases of the National Center for Biotechnology Information (NCBI), and the obtained information was used to provide detailed annotation of the phage proteome. From Table S1, it can be seen that only 65 gene products (30.4 % of the proteome) presented obvious similarity to proteins of known function and thus were tentatively assigned a corresponding function. In addition, 124 of the gene products (57.9 % of the proteome) of GH15 were found to resemble functionally unassigned proteins. The remaining 25 gene products (11.7 % of the proteome) were unique and showed no database match to any publicly available prokaryotic sequence.

When the phage gene sets were functionally classified into arbitrary categories, it was clear that the predicted genes were not randomly distributed; rather, as with phage K, these genes mapped to discrete modules (O’Flaherty et al., 2004), such as nucleotide metabolism and replication, lysis and morphogenesis, but there was no lysogeny module. Of these modules, the lysis module was the smallest one and only comprised two adjacent ORFs (ORF71 and -72). The nucleotide metabolism module consisted of two separate regions, ORF31–68 and ORF122–153. Comparative analysis indicated that there were several functional ORFs associated with the DNA replisome in this module, i.e. the DNA ligase (ORF59), RNase H (ORF63), helicase (ORF122 and -124), primase (ORF129), polymerase (ORF141) and DNA-binding proteins (ORF50 and -145) (Table S1). Furthermore, this module included ORFs encoding exonucleases (ORF125, -127 and -153), the integration host factor (ORF140) and enzymes required for DNA maturation (ORF134–136) (Table S1). In particular, ORF106 demonstrated ~30 % nucleotide identity with the putative RNA polymerase of the Enterococcus phage eEF24C. Interestingly, this ORF was located in the structural module (ORF82–117). It is worth noting that GH15 was also found to carry a putative a factor (ORF147), which encoded a protein of 26.6 kDa. This a factor could function to modify the RNA polymerase to recognize phage promoter regions, thereby regulating gene expression to express phage genes rather than host genes (O’Flaherty et al., 2004). Otherwise, both ORF47 and ORF61 encoded a 40.9 kDa and a 28.5 kDa putative ATPase, respectively. These ORFs in the nucleotide metabolism module consisted of approximately one-third of all 65 functional ORFs. Thus, phage GH15 has the potential to replicate its own DNA without too much reliance on the host. This may explain why GH15 demonstrates a very broad host range.

Phage GH15 had an overall G + C content of 30.23 mol %, which is significantly lower than that of its host, S. aureus (32.9 mol %) (Kuroda et al., 2001). The discrepancy in the G + C content between the phage and its host may result in problems of codon usage during infection (Rocha & Danchin, 2002). Coincidently, the GH15 genome was predicted to encode four tRNA genes in two regions. The first region extended from 14.8 to 15.7 kb, located in a non-coding region between ORF32 and ORF53, and contained one gene encoding Met-tRNA. The second region resembling a tRNA-encoding region was located between ORF73 and ORF74 (from 39.1 to 37.6 kb) and could encode Asp-tRNA, Phe-tRNA and a pseudo-tRNA. It had been argued that the presence of tRNAs in a phage genome compensates for differences in codon usage between the phage and its host, and that the tRNAs correspond to codons that are expected to be poorly translated by the host machinery (Bailly-Bechet et al., 2007). The second region predicted to encode tRNA genes lay adjacent to the morphogenetic cluster. The tRNA genes and their strategic locations in GH15 may enable the phage to translate its sequences more efficiently and increase its reproduction rate (Bailly-Bechet et al., 2007; Santos et al., 2011).
No restriction sites for host-encoded endonucleases

*S. aureus* is known to encode a *Sau3AI* restriction-modification system (comprising *Sau3AI, BamHI, PvuI* and *DpnI*), which recognizes the 5′-GATC-3′ sequence, and a second site-specific endonuclease, which recognizes the sequence 5′-GGNCC-3′ (O’Flaherty et al., 2004; Sussenbach et al., 1976, 1978). An interesting feature of the phage GH15 genome was that it completely lacked 5′-GATC-3′ and 5′-GGNCC-3′ (5′-GGTCC-3′, 5′-GGGCC-3′, 5′-GGACC-3′ and 5′-GGCCC3′) sequence. Hence, GH15 appears to have a very efficient mechanism of defence against these specific endonucleases. This may also be related to the broad host range of GH15.

A suspected virulence factor and a potential HNH endonuclease

Phages play an important role in the pathogenicity of *S. aureus* by carrying accessory virulence factors (Goerke et al., 2009). A BLASTP analysis revealed that the major part of GH15 ORF32 (from nt 124 to the end of the coding...
sequence, comprising 493 aa) shared 54% amino acid identity with staphylococcal lipase/acylhydrolase protein, and the G+C content of this region was 32.66 mol%, which is very similar to that of S. aureus (Kuroda et al., 2001), suggesting that this gene might be derived from its host. Alternatively, the lipase/acylhydrolase protein might be a virulence factor and associated with the lysogenic conversion of S. aureus, as described previously (Gill et al., 2005; Lee & Iandolo, 1986). However, this ORF also showed equivalent amino acid identity (45%) with the tail fibre of Staphylococcus pseudintermedius HKU10-03 phage. Thus, the real function of this gene needs further verification.

The analysis also indicated that ORF57 of GH15 showed medium sequence similarity (~30% amino acid identity) with the endonuclease HNH domain of many phages, including that of staphylococcal phage Twort (ORF52 of Twort), K (ORF31), G1 (ORF84) and others such as Salmonella phage PVP-SE1, Escherichia phage rv5 and Enterococcus phage SAP6. Thus, it can be speculated that the ORF57 of GH15 may encode an HNH endonuclease. Interestingly, GH15 ORF57 is a separate and complete gene. The corresponding homologous sequences of Twort and K/G1 are inserted into genes encoding the putative ribonucleotide reductase large-subunit protein and lysin, respectively. Additionally, the last 28 aa of GH15 ORF57 showed 82% amino acid identity with the sequence of K ORF20, which shows identity with GH15 ORF58. Thus, genetic drift and recombination could have occurred in phages.

**Comparative genomics**

Comparative analysis indicated that the identity and query coverage with GH15 escalated progressively from phage K through Sb-1, A5W, ISP and G1 to GH15 (Table 1). In addition, in the same order, the G+C content dropped significantly from 30.60 to 30.23 mol%, whilst the genome size, in contrast, progressively increased from 127 395 to 139 806 bp. As shown in Fig. 2, adjacent phages shared a high degree of homology from GH15 to G1, ISP, A5W, Sb-1 and finally K. However, two large fragments near the caudal part of A5W were absent in Sb-1 and K but existed in ISP, G1 and GH15. This finding indicated that a large-scale insertion or deletion may have occurred between these phages. It has been reported that genomes can grow (e.g. through the acquisition of ‘morons’) by non-homologous recombination, and morons sometimes have a different G+C content from that of the surrounding sequence, arguing for their recent entry into the genome (Hendrix, 2003). In light of this theory, the two large blocks of differential genes described above may be morons, and they may have been gained later in evolution based on the sequences of Sb-1 and K.

Based on the high degree of similarity, the variability of identity and coverage, the G+C content and the diverse genome size, there should be a close relationship with each other from phage K to Sb-1, A5W, ISP, G1 and finally GH15 in evolution. Thus, all of these phages could form a new group named the ‘K genus’. The G+C content of the Twort genome (30.26%) was closest to that of GH15, but, at the nucleotide level, the query coverage and identity were only 46 and 87%, respectively. It was proposed by Lavigne et al. (2008, 2009) that phages can be considered to belong to the same genus when the number of their shared homologous proteins is equal to or above 40%. According to this theory, Twort is also evolutionarily related to the K genus.

Kwan et al. (2005) described a classification system of staphylococcal phages based on their genome size: <20 kb, class I; ~40 kb, class II, and >125 kb, class III. According to this classification, in addition to the high similarity with other members of class III, GH15 can be considered a new member of class III, and all phages of this class constitute a new lineage of the family Myoviridae infecting S. aureus. These class III phages share no or very low similarity to any other S. aureus phages, which is consistent with their classification in the family Myoviridae. Additionally, for another phage, 812, its genomic size was ~146.5 kbp according to the restriction genomic map and this phage can be classified in the family Myoviridae based on its morphology (Pantuček et al., 1998). Therefore, it can be speculated that 812 also belongs to class III.

Although the overall gene organization within the genome is conserved for the majority of genes among class III phages, there are numerous insertions restricted to GH15 (Fig. S2). As described above, it has been found that two large gene fragments are completely absent in phages K, Sb-1, and Twort and are only partly present in G1, ISP and A5W. The gene fragments included nt 1050–4919 and 121 787–132 491. The G+C contents of these two fragments are 28.4 and 28.48 mol%, respectively, which is significantly lower than the G+C content of the whole GH15 genome. These mosaicisms found in these phages support the idea that large-scale genetic exchange and gene transfer have happened in prokaryotic viruses (Hendrix, 2003; Hendrix et al., 2003; Lawrence et al., 2002; Rohwer & Edwards, 2002).

**Intron-loss events in GH15**

Group I introns and inteins are intervening sequences that are autocatalytically spliced from a precursor RNA or polypeptide, respectively (Cech, 1990; Paulus, 2000). Phage-carrying group I introns typically interrupt protein-coding genes (Edgell et al., 2000). Generally speaking, phages with large genomes possess introns and/or inteins (O’Flaherty et al., 2004). The presence of group I introns has been documented previously in phages K and Twort (Kwan et al., 2005). For example, introns within the lysin (L-I) and DNA polymerase (P-I and P-II) genes of phage K are also present within their respective homologues within phages Sb-1, A5W, ISP, G1 and 812. In addition, the previously documented introns within the RNase reductase
**Table 1.** A global genome comparison of the phages analysed in this study

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>GH15</th>
<th>Twort</th>
<th>G1</th>
<th>ISP</th>
<th>A5W</th>
<th>Sb-1</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>G + C content (mol%)</td>
<td>30.23</td>
<td>30.26</td>
<td>30.39</td>
<td>30.42</td>
<td>30.47</td>
<td>30.48</td>
<td>30.60</td>
</tr>
<tr>
<td>Genome size (bp)</td>
<td>139,806</td>
<td>130,706</td>
<td>138,15</td>
<td>138,339</td>
<td>137,083</td>
<td>127,188</td>
<td>127,395</td>
</tr>
<tr>
<td>Identity of GH15 BLASTN (%)</td>
<td>100</td>
<td>87</td>
<td>97</td>
<td>97</td>
<td>97</td>
<td>97</td>
<td>97</td>
</tr>
<tr>
<td>Query coverage of GH15 BLASTN (%)</td>
<td>100</td>
<td>46</td>
<td>92</td>
<td>92</td>
<td>92</td>
<td>84</td>
<td>85</td>
</tr>
</tbody>
</table>

**Fig. 2.** A comparison of the genomic locations of homologous ORFs. Nucleotide base pairs are indicated between grey lines for each phage genome. Pairs of homologous ORFs between two phages were generated using BLASTN analysis.
represent additional examples of lateral gene-transfer content between GH15 and other homologous phages of class III. Differences in intron and intein of Twort was found (Fig. S3). This fragment is also absent comparing phages GH15 and Twort, one more inserted inteins. It is also noteworthy that, in the process of functional counterpart genes inserted with introns or inteins, GH15 shared a high degree of identity with all of these intron-negative, the corresponding gene sequences of interrupting any of the predicted genes of GH15. Despite Interestingly, screening of the complete genome sequences for relatedness between two genes of one phage (separated by one ORF) and one complete ORF of a second phage. Interestingly, screening of the complete genome sequences available indicated that there was no intron or intein interrupting any of the predicted genes of GH15. Despite being intron negative, the corresponding gene sequences of GH15 shared a high degree of identity with all of these functional counterpart genes inserted with introns or inteins. It is also noteworthy that, in the process of comparing phages GH15 and Twort, one more inserted fragment (LT-III) within the large terminase subunit gene of Twort was found (Fig. S3). This fragment is also absent in other phages of class III. Differences in intron and intein content between GH15 and other homologous phages represent additional examples of lateral gene-transfer events leading to phage mosaicism (Kwan et al., 2005).

In order to identify the presence or absence of introns/inteins, RT-PCR was performed with oligonucleotides that annealed specifically to ORF71, ORF82, ORF100, ORF122, ORF135 and ORF141 of GH15. RNA was isolated from S. aureus strain W3275 10 min after infection by GH15 and used for the synthesis of cDNA. DNA samples were amplified with the corresponding primers. As controls, GH15 DNA and RNA isolated from GH15-infected cells were PCR amplified with the same primers. From Fig. 3, it can be seen that these ORFs at the level of the genome were the same size with those at the cDNA level. This result indicated that RNA splicing had not occurred in these six ORFs and was consistent with the results of genome analysis.

The introns that exist in the lysin and polymerase genes of most class III phages do not share homology with any gene in other phage or bacterial genome. The unique nature of these introns may support the speculation that these introns originated in the phages and that there have been intron-loss events in GH15. Aside from this possibility, this conclusion is in agreement with the theoretical expectation that selective pressures would remove the intron (Darnell & Doolittle, 1986). Our analysis suggested that most of the observed intron deletions are exact. In addition, most of these intron- and intein-related genes are highly expressed. All of these features support the previous hypothesis that intron loss is mediated via the homologous recombination of intronless cDNA of the corresponding gene (Coulombe-Huntington & Majewski, 2007). Therefore, there may have been intron-loss events in GH15 instead of the intron-gain events documented in other phages.

**Comparison of sequences surrounding splice sites**

The DNA sequences of a 40 bp region surrounding the intron insertion site of related genes were compared with the equivalent regions of the corresponding intron-free genes (Fig. 4). Because of the uncertainty of the insertion site of intron RR-I, the comparison of this intron could not be performed. The sequence alignment demonstrated that there were clear differences between intron-containing and intron-free alleles in a 20 bp segment centred around the intron insertion point. Group I introns can encode homing endonucleases, which are responsible for placing the intron in the correct sequence context for efficient splicing (Foley et al., 2000). If there are mutations in the specific sites recognized by the corresponding endonuclease, this may bring about the failure of intron insertion (Chevalier & Stoddard, 2001; Haugen et al., 2005). The differences in sequences near the inserted sites between intron-containing and intron-free alleles most likely lead to the establishment of the consensus sequences associated with the presence of introns. It also reflects the critical nature of these nucleotide positions for intron homing and splicing (Foley et al., 2000).

**Proteome analysis of the GH15 particles**

To characterize phage GH15 further, the structural protein composition was analysed by SDS-PAGE and mass spectrometry. Following gel electrophoresis, 13 major protein bands were visualized by Coomassie Brilliant Blue staining and analysed by mass spectrometry (Fig. 5).
Fig. 4. Nucleotide sequence alignment of the region surrounding the splice site of introns or inteins in ORF71 (lysin) (a), ORF141 (polymerase) (b), ORF82 (large terminase) (c), ORF100 (capsid) (d), ORF122 (DNA helicase) (e) and ORF135 (ribonucleotide reductase) (f). Sequence differences are indicated using different shading. A dashed line indicates intron-free or intein-free regions. Intron and intein insertion sites are indicated by vertical arrows.
Ten structural proteins were ultimately identified. The major capsid protein (ORF92) was most probably present in the most prominent protein band 6. Band 12 was also a putative capsid protein (ORF100). The tail of GH15 is complex and five tail-related proteins were identified (bands 1, 4, 7, 11 and 13). According to the mass spectrometry analysis and molecular masses, bands 1, 4, 7, 11 and 13 contained ORF117 (putative adsorption-associated tail protein), ORF99 (putative major tail sheath protein), ORF114 (showing 56% identity with the Listeria phage A511 gp103 putative tail protein) and ORF149 and ORF150 (putative major tail proteins), respectively. It is worth noting that band 3 also contained ORF99, and may be a modified tail sheath protein. In addition, bands 2, 5 and 9 contained ORF110 (showing 98% identity with a putative glycrophosphoryl diester phosphodiesterase), ORF89 (portal protein) and ORF148 (showing 34% identity with Enterococcus phage ϕEF24C putative Ig-like protein), respectively. In proteins bands 8 and 10, no structural protein with a corresponding molecular mass was found. Most of these identified structural proteins were located in the structural module, with the exception of ORF148–150, which were located in the large nucleotide metabolism and replication module.

Interestingly, ORF89, -92, -99, -110, -114, -117 and -148 of phage GH15 showed homology with the corresponding structural proteins of the Enterococcus phage ϕEF24C-P2 and Listeria phages P100 and A511 (Lavigne et al., 2009). Indeed, there were a number of other GH15 proteins that were found to be homologous with the predicted genes of these three phages. Some of these proteins were distributed in the nucleotide metabolism and replication module [such as the putative large terminase subunit (ORF82), RNA polymerase (ORF106), helicase (ORF122), transcriptional regulator (ORF123), exonuclease (ORF125), DNA repair exonuclease (ORF127), primase (ORF129), resolvase (ORF133) and DNA polymerase (ORF141)] (Table S1). All these proteins are essential for phage survival and belong to the set of ‘core proteins’ of the phage (Krupovic et al., 2011). The similarity suggested that GH15 and these three phages (in other species) are related and may have descended from a common ancestor.

In conclusion, we have described here the genome sequence of a novel staphylococcal myovirus phage, GH15. Detailed genetic characterization of this phage revealed a number of features: (i) the genome sequence of GH15 showed high similarity to most of the phages in class III and there were close evolutionary relationships from K to Sb-1, A5W, ISP, G1 and finally to GH15; (ii) there were no introns or inteins interrupting genes with crucial enzymic functions in the GH15 genome; (iii) there may have been intron-loss events in GH15 instead of the intron-gain events documented in other phages; (iv) there were a large number of mosaicisms shared among GH15 and other phages; and (v) most of the ‘core proteins’ of GH15 showed similarity to corresponding proteins of Enterococcus phage ϕEF24C-P2 and Listeria phages P100 and A511. The comparative analysis of the GH15 genome sequence with other phages not only provides compelling evidence for the diversity of the staphylococcal myovirus phages but also offers new clues to intron loss in phages.

**METHODS**

**Source of phage and bacterial strain.** The phage GH15 used in this study was isolated from the Changchun (China) sewage system, and the phage host, meticillin-resistant *S. aureus* W3275, was obtained from the First Hospital of Jilin University (Changchun, China) (Gu et al., 2011b) and routinely subcultured in nutrient broth with supplements (per litre: 10 g nutrient broth, 5 g peptone, 2.5 g yeast extract, 9 g NaCl, 0.3 g KCl) with shaking at 37 °C (Landthaler & Shub, 1999).

**Phage propagation.** *S. aureus* cells were grown to an optical density at 600 nm of 0.2 and infected at an m.o.i. of 0.1 phage per cell; the incubation was continued until lysis was complete. The lysate was centrifuged at 4 °C (10,000 g; 15 min) and the supernatant was filtered through a 0.22 μm pore size Millipore filters. Concentrated phage GH15 preparations were obtained by CsCl density-gradient centrifugation following a 10% polyethylene glycol 8000 precipitation of phage-containing lysates as described by Sambrook and Russell (2006). The band with the highest opalescence was collected and dialysed against SM buffer [50 mM Tris/HCl (pH 7.5), 100 mM NaCl, 8 mM MgSO₂] in a Slide-A-Lyser cassette with a molecular-mass cut-off of 10,000 Da (Pierce Biotechnology).
A sample was applied to copper grids, negatively stained with 2% (w/v) phosphotungstic acid and examined by transmission electron microscopy (JEOL JEM-1200EXII; Japan Electronics and Optics Laboratory) at an accelerating voltage of 80 kV.

Bioinformatic analyses. Sequence analysis was performed using DNASTAR software. The circle map of the GH15 genome was made with CGView (Grant & Stothard, 2008). Coding sequences were predicted and analysed by BLAST (Altschul et al., 1990) and GeneMarkS (Besemer et al., 2001). Genes encoding tRNA and rRNA were identified by tRNAscan-SE (Lowe & Eddy, 1997) and RNAmmer (Lagesen et al., 2007), respectively. The gene and deduced amino acid sequences were used for database searches (using standard parameters) with BLASTN and BLASTP, respectively (provided by the NCBI network service) (Altschul et al., 1997). Global genome comparisons were constructed using the BLASTN Artemis Comparison Tool (Carver et al., 2008).

RT-PCR. The phage GH15 DNA was extracted from a CsCl-purified phage stock solution (Kwan et al., 2005). RNA isolation and RT-PCR were carried out as described previously with some modifications (Landthaler & Shub, 1999). In short, RNA isolated from S. aureus W2375 10 min after infection with GH15 was reverse transcribed using RevertAid Moloney murine leukemia virus reverse transcriptase (MBI Fermentas) and oligo(dT)18, according to the manufacturer’s protocol. DNA samples were amplified in a Biometra T3000 thermocycler programmed for 35 cycles each consisting of 94°C for 45 s, 48°C for 45 s and 68°C for 2.5 min, using KOD DNA polymerase, according to the manufacturer’s protocol. The sequences of the relevant primer pairs used for PCR and RT-PCR were as follows: ORF71P-F (5'-ATGCTTGAAGTATTTCA-3') and ORF71P-R (5'-CTATTTGAATACTCCCCAGGCAA-3'), ORF82P-F (5'-AAAGAATATTAGATGCACGAAGA-3') and ORF82P-R (5'-TTCCATGAAAAGGATATATTGTTTATTT-3'), ORF100P-F (5'-TTAGGTGTTTTCGGATCAGTCAACAGCAAACGGCCG-3') and ORF100P-R (5'-TTAGGTGCCTTGGTCTCACCTGCT-3'), ORF112P-F (5'-ATGGAGATTAGAATTAAGACATTAT-3') and ORF112P-R (5'-CTATTTTCCTATTATTGTTTAAAGA-3'), and ORF141P-F (5'-ATGATAGTATATTTTCTTGG-3') and ORF141P-R (5'-TCTAATTTCTGTAAATATGT-3').

SDS-PAGE and proteomic analysis. The structural proteins of GH15 were analysed by SDS-PAGE and mass spectrometry (Sanutos et al., 2011). The concentrated phage samples were mixed with the sample buffer (containing 2 mM 2-mercaptoethanol), heat denatured (100°C, 8 min) and loaded onto a standard SDS-polyacrylamide gel (12% acrylamide) for separation of proteins. The gels were stained with Coomassie Brilliant Blue R250 and the entire lane was cut into slices and subjected to trypsin digestion (Shevchenko et al., 1996). The peptides generated were subsequently analysed by electrospray ionization tandem mass spectrometry (ESI-MS/MS) as described previously (Lavigne et al., 2006). The peptide masses and fragmentation spectra obtained were analysed using BLASTP and against the genome of phage GH15.

GenBank accession numbers. The GenBank accession numbers of phages GH15, K, Sb-1, ASW, ISP, G1 and Twort used in this study are JQ686190, AY176327, HQ163896, EU418428, FR852584, AY954969 and AY954970, respectively.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Key Program, no. 311130072).

REFERENCES


GH15 provides new clues to intron shift


