JOURNAL OF BACTERIOLOGY, Apr. 2007, p. 2646–2652 Vol. 189, No. 7
0021-9193/07/$08.00+0 doi:10.1128/JB.01590-06

Phage 3396 from a *Streptococcus dysgalactiae* subsp. *equisimilis* Pathovar May Have Its Origins in *Streptococcus pyogenes*▽†

Mark R. Davies,[1,2,3] David J. McMillan,[1] Gary H. Van Domselaar,[4,5] Malcolm K. Jones,[6] and Kadaba S. Sriprakash[1,2,3]*

Bacterial Pathogenesis Laboratory, The Queensland Institute of Medical Research, Brisbane 4029, Australia[1]; The Australian Centre for International Tropical Health and Nutrition, The Queensland Institute of Medical Research, Brisbane 4029, Australia[2]; Cooperative Research Centre for Vaccine Technology, The Queensland Institute of Medical Research, Brisbane 4029, Australia[3]; National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, MB R3E 3R2 Canada[4]; Faculty of Pharmacy, University of Manitoba, Winnipeg, MB R3T 2N2 Canada[5]; and Molecular Parasitology Laboratory, The Queensland Institute of Medical Research, Brisbane 4029, Australia[6]

Received 13 October 2006/Accepted 17 January 2007

***Streptococcus dysgalactiae* subsp. *equisimilis* strains (group G streptococcus [GGS]) are largely defined as commensal organisms, which are closely related to the well-defined human pathogen, the group A streptococcus (GAS). While lateral gene transfers are emerging as a common theme in these species, little is known about the mechanisms and role of these transfers and their effect on the population structure of streptococci in nature. It is now becoming evident that bacteriophages are major contributors to the genotypic diversity of GAS and, consequently, are pivotal to the GAS strain structure. Furthermore, bacteriophages are strongly associated with altering the pathogenic potential of GAS. In contrast, little is know about phages from GGS and their role in the population dynamics of GGS. In this study we report the first complete genome sequence of a GGS phage, Φ3396. Exhibiting high homology to the GAS phage Φ315.1, the chimeric nature of Φ3396 is unraveled to reveal evidence of extensive ongoing genetic diversity and dissemination of streptococcal phages in nature. Furthermore, we expand on our recent findings to identify inducible Φ3396 homologues in GAS from a region of endemicity for GAS and GGS infection. Together, these findings provide new insights into not only the population structure of GGS but also the overall population structure of the streptococcal genus and the emergence of pathogenic variants.**

Streptococcus dysgalactiae subsp. *equisimilis* (group G streptococcus [GGS]) commonly inhabits the throat, skin, and vagina of healthy humans (28). However, occasionally the organism can cause pharyngitis, impetigo, cellulitis, septicemia, glomerulonephritis, and toxic shock (14, 28). This same disease spectrum generally coincides with that of a well-known human pathogen, the group A streptococcus (GAS; *Streptococcus pyogenes*) which cohabits the same tissue sites. Due to the extensive overlap of the disease spectrum of GGS with that caused by GAS, it is possible that disease burden specifically attributed to GGS has been historically underestimated. Of late, there are reported increases of GGS infections associated with classic human GAS-associated diseases, including invasive manifestations (10, 14, 15, 19, 23, 34).

Phylogenetic analyses based on rRNA sequence and other molecular clocks revealed that *S. dysgalactiae* spp. are closely related to *S. pyogenes* (16, 26, 35, 38). The observations of overlapping clinical presentations, niche-sharing for colonization, and this evolutionary relatedness are conducive for interspecies lateral genetic transfers (LGT). Previous studies have shown that GGS possesses genes for M protein, C5a peptidase, streptokinase, streptococcal pyrogenic exotoxins, and fibronectin binding proteins (9, 20, 21, 24, 25, 29, 37, 40). These are well defined as essential virulence determinants in GAS. The presence of mosaic structures among some of these genes within GAS and GGS strongly suggests that interspecies LGT events do indeed occur between these two streptococci and could be derived from recent and ongoing events, as demonstrated in our recent study (13). Most of these LGTs are likely to be mediated by phages and other mobile genetic elements.

Evidence is accumulating that phages contribute enormously to strain structure in GAS (2, 17). So far, six major families of phages have been reported in GAS (7). Numerous virulence characteristics in this pathogen are phage associated, including the pyrogenic exotoxins, DNase, and mitogenic factors (12). Our recent studies (D. J. McMillan, R. Geffers, J. Buer, B. J. M. Vlaminckx, K. S. Sriprakash, and G. S. Chhatwal, submitted for publication) using a targeted microarray representing genes for GAS virulence factors and extracellular proteins revealed that differences in the genetic profiles of these genes in GAS are largely restricted to 11 genomic loci, and most of these have the footprints of mobile genetic elements in close proximity.

By contrast, little is known about phages from GGS or about the role of phages in the dynamics of GGS population structure in nature. We have shown that phages are by far the major contributors to the differences seen between a GGS pathogenic isolate (GGS pathovar) and a commensal isolate from a

* Corresponding author. Mailing address: Bacterial Pathogenesis Laboratory, Queensland Institute of Medical Research, Brisbane, Queensland 4006, Australia. Phone: 61 7 33620407. Fax: 61 7 38453507. E-mail: sriS@qimr.edu.au.

† Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 26 January 2007.

healthy individual (13). Furthermore, these phages appeared to be related to families of known GAS phages. These observations raise important questions in regard to the influence of circulating GAS strains, particularly where the GAS disease burden is high, on the genesis of "pathovaric" characteristics within commensal GGS.

Here, we report the first complete sequence of a GGS phage, Φ3396, recovered from a GGS pathovar. The results show that this phage is phylogenetically linked to one of the GAS phage families (315.1-like phage) but exhibits evidence of modular recombination in the DNA replication and lysogenic conversion regions from other streptococcal phages. Generation of such chimeric phages could expand the repertoire of virulence and other genes that could be mobilized from pathogenic to the commensal bacteria. Expanding on our recent findings (13), inducible Φ3396 homologues were identified only within GAS from a area of endemicity for GAS infection, consistent with the interpretation that these chimeric phages may have arisen in GAS and subsequently transferred to GGS in an environment where GAS is endemic.

MATERIALS AND METHODS

Bacterial strains and growth conditions. Phage Φ3396 was isolated from a GGS pathovar isolate (NS3396) obtained from a 15-year-old Aboriginal male who presented with recurrent acute rheumatic fever (13). The distribution of Φ3396-related phage was examined in 10 GGS sterile site (blood culture) isolates from bacteremic patients in Brisbane, Australia, and 11 GGS isolates from a mixture of invasive and colonizing strains from the Northern Territory (NT) of Australia, where GAS infections are endemic. A further 12 GAS isolates from Sydney, Australia, and 20 GAS isolates from the NT were also screened for the presence of a phage Φ3396 homologue. All stock cultures for long-term storage were under identical conditions, namely, glycerol cultures at −70°C. Standard nonparametric chi-square analysis was performed to examine statistical correlations between the distribution of the phage and phage elements. All isolates were cultured on Todd-Hewitt agar (Oxoid, England) supplemented with 2% horse blood or in Todd-Hewitt broth at 37°C.

DNA extractions and *emm* sequencing. Genomic DNA was extracted using a QIAGEN DNeasy kit, with the addition of mutanolysin (285 units/ml) (Sigma) to enhance cell wall degradation. The presence of the *emm* gene was determined by PCR using the primers 5′-CAGTATTCGCTTAGAAAATTAAAA-3′ and 5′-CAGAGCTGTTGCCATAACAGTAAG-3′ from the leader and carboxy regions of the gene, respectively. The *emm* type was determined by BLAST analysis of the *emm* sequence database (www.cdc.gov/ncidod/biotech/strep/strepindex.htm). DNA sequencing was performed using an ABI Big Dye 3.1 dye terminator sequencing kit in conjunction with an ABI 3100 Genetic Analyser (Applied Biosystems).

Phage modular screening and identification of flanking regions by adaptor PCR. The presence of the Φ3396 prophage within GAS and GGS isolates was determined by screening for three Φ3396 regions by PCR. The three regions amplify genes within five different phage modules: lysogeny ([region one] Φ3396_0682F, 5′-TCTTATCCAAATCGGAACGG-3′; Φ3396_0682R, 5′-TTTTCTGCCCCTCACTATGC-3′); DNA packaging, head-joining-tail, and tail modules ([region two] Φ3396_0713F, 5′-GCGTCTAAAAGCTCTTACCG-3′; Φ3396_0721R, 5′-CTAACAAATCTTGCCAATCAG-3′); and a region within the lysogenic conversion region ([region three] Φ3396_0734F, 5′-AAATACCCACAACGTTTTTCAGC-3′; Φ3396_0734R, 5′-AATTTACAATAAACCTTTTAAATATCG-3′). Standard conditions for all PCRs included an initial denaturing step of 2 min at 95°C followed by 30 cycles of 95°C for 15 s, 50°C for 30 s, and 72°C for 2 min.

Identification of the Φ3396 prophage flanking genes was determined in NS3396 and the GAS isolates NS14, NS32, NS50.1, NS80, and NS195 through adaptor PCR. Adaptor PCR was performed as previously described (27) with minor modifications. Essentially, genomic DNA was partially digested with BclI and HindIII restriction endonucleases (New England Biolabs) and one of two 5′ phosphorylated oligonucleotide adaptors containing BclI and HindIII overhangs (5′-GATCCGCCTATAGTGAGTCGTATTAAC-3′ or 5′-AGCTCGCCTATAGTGAGTCGTATTAAC-3′) ligated to the digested DNA. PCR using prophage-specific primers corresponding to the left proximal gene (Φ3396_0682R3, 5′-TTTCGTGTTACCCCTCACATGATGC-3′) or the right proximal gene (Φ3396_0734F) in conjunction with the adaptor-specific primer (5′-GTTAATACGACTCACTATAGGCG-3′) was performed to amplify into the genomic regions flanking the Φ3396-like prophages. Products were subsequently purified by agarose gel electrophoresis (QIAGEN), and the DNA sequences were determined.

Phage induction and purification. Isolates that were positive by PCR for the presence of all Φ3396 prophage regions were treated with mitomycin C to test for the presence of inducible phages as previously described (3). Briefly, bacteria were cultured overnight at 37°C in 2× Todd-Hewitt broth. The overnight culture was diluted 1:50 into prewarmed (held at 37°C) Todd-Hewitt broth and grown to an optical density at 600 nm of 0.2 before induction with 0.2 μg/ml of mitomycin C (Sigma). Fifty-microliter aliquots were taken at preinduction (0 min) and at 180 min postinduction. The aliquots were pelleted at 16,000 × *g* for 5 min, and cells were lysed using colony lysis procedures (18). To test for the excision of the prophage from the bacterial chromosome, PCRs were performed using divergent primer pairs at opposing ends of the prophage sequence (Φ3396_0734F and Φ3396_0682R). Only when the phage DNA was excised and circularized from the host genome would the primers produce a PCR product across the phage attachment (*att*) sites. To allow for heterogeneity within the lysogenic conversion region, a second set of divergent primers was used (Φ3396_0734F2, 5′-GGATACCAATCAGCAAACAATCGGC-3′, and Φ3396_0682R). Identification of *attB*- and *attP*-proximal sequences was carried out by ClustalW alignment of sequences from induced prophage and prophage-flanking regions.

Electron microscopy. Isolate NS3396 was treated with mitomycin C (as described above) for 4 h. Cellular debris was pelleted at 8,000 × *g* for 15 min. The induced supernatant was filtered through a 0.44-μm-pore-size filter (Millipore) and centrifuged at 100,000 × *g* for 2 h at 4°C. Pelleted phage particles were resuspended in 500 μl of SM buffer (5g liter^{-1} NaCl, 2g liter^{-1} $MgSO_4$, 50 mM Tris-HCl [pH 7.5]) and negatively stained with uranyl acetate (3). Phage preparations were examined using a JEM 1010 transmission electron microscope operated at 80 kV. Digital images were recorded using an AnalySIS Megaview III digital camera.

Phage sequencing, assembly, and comparative genomics. Long PCR and primer walking techniques were used to obtain overlapping fragments of the Φ3396 prophage. Primer selection was based on Φ3396 gene fragments identified through genomic subtraction of NS3396 (13) in conjunction with primers designed from the MGAS315.1 phage (Φ315.1) homologue (NC_004584) (4). Forward and reverse strand contigs were assembled using the Staden package (5). Annotation of phage Φ3396 was performed using the BASys bacterial annotation system (39). Whole-phage genomic comparisons of Φ3396 and the GAS phages Φ315.1 and Φ8232.1 were performed using the Artemis comparison tool (8). Dot plots were performed using Dotter (31).

Nucleotide sequence accession number. The Φ3396 sequence was deposited in the GenBank database under accession number EF207558.

RESULTS

General features of phage Φ3396 and its genome. Through a process of long PCR and primer-walking methods, the complete genome sequence of Φ3396 was determined. The genome is 38,528 bp with an average GC content of 37.48%. Annotation of the Φ3396 genome was performed using the bacterial annotation system BASys (39). A total of 64 open reading frames (ORFs) were predicted (see Table S1 in the supplemental material), with an average translated ORF of 188 amino acids (aa). GenBank database analysis revealed that 10 ORFs had little or no homology to entries within the protein database. Such ORFs were generally small, averaging 246 bp. A further 30 ORFs (47%) were homologous to hypothetical phage proteins (average of 432 bp), with the remaining ORFs corresponding to known phage-associated genes. The overall genetic organization of Φ3396 is consistent with that reported for all currently known phages of lactic acid bacteria (7): left attachment site, lysogeny, DNA replication, transcriptional regulation, DNA packaging, head-joining-tail-tail fiber, lysis, and right attachment site (Fig. 1).

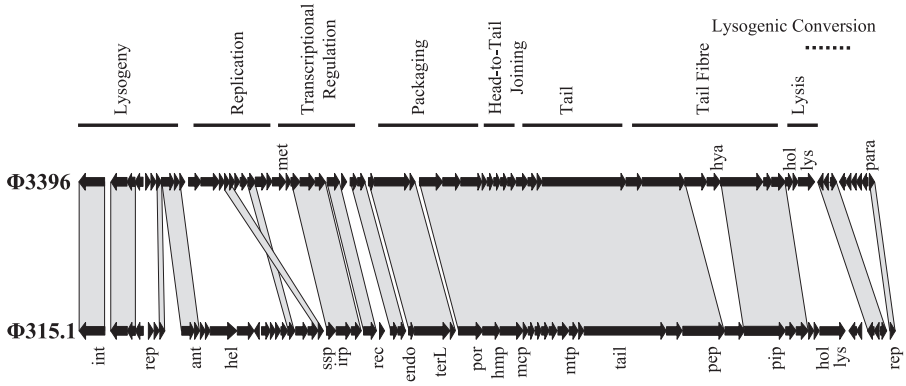

FIG. 1. Genetic organization and protein alignment of two phage sequences: *S. dysgalactiae* subsp. *equisimilis* phage Φ3396 and *S. pyogenes* phage Φ315.1 (4). Conserved proteins exhibiting greater than 50% amino acid identity are represented by the shading. Genomes were annotated using the bacterial annotation system BASys (39) with comparisons performed using the Artemis comparison tool (8). The Φ315.1 genome sequence was obtained from the GenBank database (NC_004584) (4). The prophage modules, as identified by bioinformatics analyses, are represented above the figure. Selected proteins and genes are defined and represented in alphabetical order as follows: *ant*, antirepressor; *endo*, endonuclease; *hel*, helicase; *hol*, holing; *hmp*, head maturation protease; *hya*, hyaluronoglucosaminidase; *int*, integrase; *irp*, immunity repressor protein; *lys*, hydrolase/lysin; *mcp*, major coat protein; *met*, DNA methylase; *mtp*, major tail protein; para, paratox; *pep*, endopeptidase; *pip*, phage infection protein; *por*, portal protein; *rec*, recombinase; *rep*, repressor; *ssp*, single-strand binding protein; tail, tail protein; *terL*, large subunit terminase.

Of the 54 Φ3396 translated ORFs with greater than 35% identity to proteins in the nonredundant GenBank database, 47 (87%) were homologous to GAS bacteriophages; 37 (79%) of these exhibited greater than 80% amino acid identity (see Table S1 in the supplemental material). Another 5 of the 54 translated ORFs were homologous to phage-like genes from the group B streptococcus (GBS) ΦSA1 ($n = 1$), ΦSA2 ($n = 3$), and GBS TIGR01671 ($n = 1$); however, four of these five ORFs had less than 60% amino acid identity. The remaining two encoded proteins exhibit less than 50% amino acid identity to phage genes from *Flavobacterium* sp. and *Enterococcus faecalis* and may represent highly divergent homologues.

Comparative genomics of Φ3396 and GAS phages. From the above analysis, it was apparent that Φ3396 exhibited significant homology to characterized GAS phages, particularly the M3 GAS phage 315.1 (NC_004584) (4). Of the 47 GAS homologues in Φ3396, 34 (72%) were homologous and syntenic to Φ315.1. An additional two ORFs were homologous to the M18 GAS phage Φ8232.1 (NC_003485.1) (30), which is also a member of the 315.1 phage family (7). As in the 315.1 phage family, the highly conserved gene paratox (1) was also present in Φ3396. Further analysis of the 37 homologues in the 315.1 phage family revealed that a total of 34 (92%) share greater than 90% amino acid identity at the protein level. Given that these homologues occur through the whole length of the Φ3396 genome (Fig. 1), this phage is most likely a member of the 315.1-like GAS phage family.

While the degree of genetic conservation across the Φ3396 and Φ315.1 genomes is generally high, four regions of Φ3396 exhibited little or no similarity with this family (Fig. 2). The first region encompassing Φ3396 ORF3 to ORF7 lies within the lysogeny module and shows high levels of nucleotide diversity with Φ315.1 while exhibiting similarity to two GBS ΦSA2 genes (36) and the 315.1-like phage Φ8232.1 (30). The second region (Φ3396 ORF12 to ORF22) encompasses the

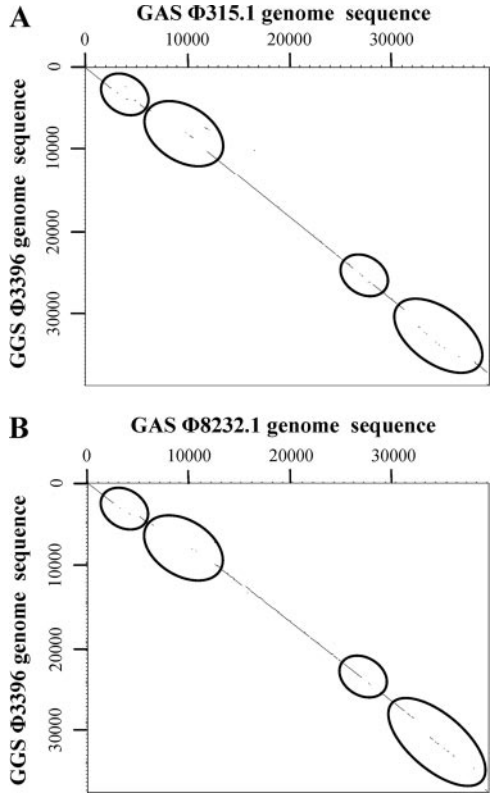

FIG. 2. Dot plot comparisons of the DNA genome sequences of the GGS bacteriophage Φ3396 (vertical axis) and the GAS prophages Φ315.1 (A) and Φ8232.1 (B). Regions of high DNA homology are represented with regions of significant nucleotide differences indicated by circles. Relative base locations are indicated on the axes. Dot plots were performed using Dotter (31).

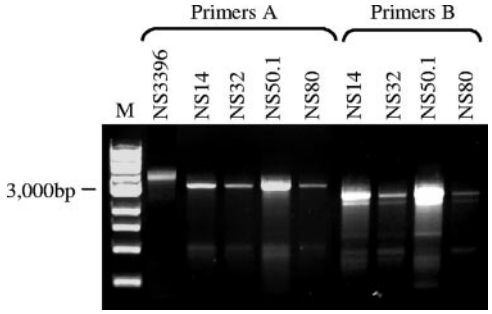

FIG. 3. PCRs showing mitomycin-induced excision of the 315.1-like prophage from the genome of GGS NS3396 and four NT GAS isolates: NS14, NS32, NS50.1, and NS80. PCRs were performed using two sets of divergent primer pairs at opposing ends of the prophage sequence (Φ3396_0734F and Φ3396_0682R [primers A]; Φ3396_0734F2 and Φ3396_0682R [primers B]). Only when the phage DNA was excised from host DNA and formed a circular intermediate did the primers produce a product across the phage attachment site. The differences in size between the NS3396-induced product and the GAS isolates may represent recombination events at the distal ends of the phage. M is the 1-kb DNA marker (New England Biolabs).

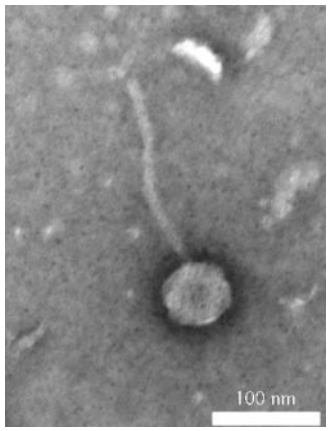

FIG. 4. Electron microscopy of phage Φ3396 negatively stained with uranyl acetate. Bar, 100 nm.

entire DNA replication module (Fig. 1). This region primarily retains low homology to genes from various non-315.1-like GAS phage families and GBS phage genes, suggesting potential modular recombination events. The DNA methylase gene has low homology to *Flavobacterium* sp. and may represent a unique methylase. There is also evidence of a gene duplication involving ORF16 and ORF23 encoding the hypothetical gene SpyM3_0700. The third region of low homology lies within the Φ3396 tail protein (ORF45). ClustalW alignment at the protein level between Φ3396 ORF45 and its Φ315.1 homologue, SpyM3_0719, showed an overall 72% amino acid identity over the 1,375-aa protein. While the first 639 aa of the phage tail protein showed 92% protein identity, the remaining 736 aa exhibited higher degrees of variation (56% identity) and may constitute a level of evolutionary adaptation of Φ3396 for the GGS cell wall. The fourth region of low homology encompasses the tail fiber, host lysis, and lysogenic conversion modules, indicating that Φ3396 may have evolved a unique set of GGS-adapted proteins for attachment to the bacterial surface, as well as lysogeny and lysis. However, heterogeneity within the lysogenic conversion region is common within streptococcal phages including the 315.1 phage family (7).

Distribution of Φ3396 phage modules in GAS and GGS isolates. From our previous studies, we suggested that Φ3396 was detected frequently in GGS isolates from the NT, Australia, an area where GAS infections are endemic, as opposed to isolates from an area where these infections are not endemic (13). To further elaborate on these findings, PCR of three Φ3396 regions encompassing five phage modules (lysogeny [region one]; DNA packaging, head-joining-tail, and tail modules [region two]; and lysogenic conversion [region three]) was used to specifically screen for this phage among nonclonal GGS ($n = 11$) and GAS ($n = 20$) isolates from an area (NT) where GAS is endemic and contrasted its distribution among nonclonal GGS ($n = 10$) and GAS ($n = 12$) isolates from an area where GAS is nonendemic (see Table S2 in the supplemental material). By *emm* sequence type, geographical and temporal differences of the isolates, phage markers (for all, see Table S2 in the supplemental material), and other chromosomal markers as determined by virulence microarray (unpublished data), we consider that our collection of strains represent predominantly nonclonal isolates. Of the 21 GGS isolates screened, only NS3396 was positive for all the Φ3396 regions tested. An additional 2 of the 11 (combined total of 27%) GGS isolates from the NT were positive for two regions. By contrast, no GGS strains from an area where GAS was nonendemic possessed any Φ3396 regions. Seven of the 20 (35%) NT GAS isolates were positive for a minimum of two phage regions, of which four were positive for all three regions. Given that these isolates represented seven different *emm* types, it is unlikely that they represent a clonal lineage (see Table S2 in the supplemental material). Furthermore, none of the GAS isolates from an area where GAS is not endemic were positive for any Φ3396 regions. Together, these results clearly show that Φ3396-like phages were statistically ($P < 0.05$) more frequently present in streptococcal isolates from areas of GAS endemicity than in areas of GAS nonendemicity. However, given the small sample sizes, no statistical difference between the frequencies of Φ3396 in GAS or GGS isolates could be identified.

Inducibility of prophages related to Φ3396 from the GAS strains identified to possess the prophage. Previously, we demonstrated that Φ3396 was inducible from isolate NS3396 after mitomycin treatment by using divergent primers at the ends of the prophage (13). Circular replicative forms of induced phage would bring the primers into the correct orientation to yield an approximate 4-kb product. Similar tests were performed on the four GAS strains (NS14, NS32, NS50.1, and NS80) that contained all three Φ3396 phage regions (Fig. 3; see Table S2 in the supplemental material). To account for heterogeneity within the lysogenic conversion region, two primer combinations were used. All four GAS strains yielded a smaller PCR product after induction than NS3396 (Fig. 3).

Electron microscopy of mitomycin-induced and purified Φ3396 revealed an icosahedral capsid with a long noncontractile tail of approximately 180 nm (Fig. 4). In addition to the genomic organization of Φ3396, these observations are consis-

TABLE 1. PCR screening of phage attachment sites within Φ3396-positive GAS isolates[a]

Isolate	PCR result with the indicated primer pair			
	Genomic flanking lysogeny[c]	Φ3396 flanking lysogeny (variation)[d]	Genomic flanking lysogenic conversion[c]	Φ3396 flanking lysogenic conversion (variation)[d]
	SpyM3_0680F and SpyM3_0680R	Φ3396_0682R3 and SpyM3_0680F	SpyM3_0737F and SpyM3_0737R	Φ3396_0734F and SpyM3_0737R
NS3396[b]	+	+	+	+
NS14	+	+	+	+ (2 kb smaller)
NS32	+	+ (300 bp larger)	+	+ (2 kb smaller)
NS50.1	+	+ (200 bp larger)	+	+ (2 kb smaller)
NS80	−	+ (500 bp larger)	+	+

[a] GAS isolates positive for all three phage regions (see Table S2 in the supplemental material) were termed Φ3396 positive. +, presence of a PCR product of expected size unless otherwise stated; −, no PCR product detected. Size variations were approximated from agarose gel analysis of PCR products and compared to the NS3396 control; PCR primers for the flanking genes SpyM3_0680 and SpyM3_0737 PCR were designed from the GAS genome sequence MGAS315 (NC_004070.1) (4).

[b] NS3396 is the GGS isolate from which Φ3396 was isolated.

[c] Genomic flanking: primer combinations targeting genes neighboring the phage/bacterial attachment sites at either the lysogeny or lysogenic conversion ends of Φ3396.

[d] Φ3396 flanking: primers targeting both phage and genomic genes across the phage/bacterial attachment sites.

tent with the *Siphoviridae* family of phages common to streptococci.

Identification of the genes flanking prophage 3396. PCR fragments corresponding to genes flanking the prophage were amplified using an adaptor-specific primer and primers targeting the distal ends of the prophage. Bcl1-digested NS3396 DNA ligated to the adaptor gave a 2-kb PCR product with Φ3396_0682R3 and the adaptor-specific primer. BLAST analysis revealed that the left flanking gene in NS3396 was highly homologous to the hypothetical protein SpyM3_0680 from the M3 GAS 315 genome. Interestingly, this gene is known to lie adjacent to the left attachment site of prophage 315.1. Similarly, adaptor PCR using primer Φ3396_0734F specific to the lysogenic conversion region and the adaptor-specific primer produced a 5-kb PCR product which is highly homologous to the GTP-binding protein LepA (SpyM3_0737) in GAS 315. This gene also harbors the right attachment site of Φ315.1. Collectively, these results show that both the Φ3396 and Φ315.1 lysogenize their respective hosts by integrating at similar genomic localizations.

In order to examine the insertion sites of the 315.1-like phage in GAS from the NT, where GAS is endemic, PCR was conducted on the four NT GAS isolates that were found to harbor all three Φ3396 phage regions (Table 1). All four isolates tested gave a PCR product across the lysogeny module to the hypothetical flanking gene SpyM3_0680. However, size variation in the amplified product may indicate sequence variation between isolates (Table 1). Similarly, all four isolates were PCR positive across the 315.1-like right attachment site (SpyM3_0737). One isolate, NS80, gave a 5-kb product, similar to NS3396, while three isolates gave a smaller 3-kb product. Sequencing of the smaller 3-kb product from NS14 revealed a 2-kb difference between the SpyM3_0736 gene and the paratox gene, similar to the sequence found in Φ315.1.

DISCUSSION

It is becoming increasingly evident that phages play an important role in the intraspecies genetic variation of many bacterial species (6). A prime example is GAS, whose genome is often polylysogenic, with phages accounting for up to 75% of the intraspecies genetic differences (2, 7, 17). However, the extent of their role in the emergence of genetic variations among other pathogenic members of the genus *Streptococcus* is less well known. While in vitro interspecies transduction experiments were documented in streptococci in the 1970s (11, 33), it is not known whether such transfers indeed occur in nature.

Comparatively, the role of bacteriophages in the genetic makeup of GGS and their effect on streptococcal population structure are yet to be elucidated. In this study we have isolated and sequenced the first full-length inducible phage (Φ3396) from *S. dysgalactiae* subsp. *equisimilis*. Genetically and morphologically, Φ3396 belongs to the *Siphoviridae* family of phages, of which all GAS phages known to date are members (7). Comparative genomic analysis revealed that Φ3396 exhibits an overall identity of 73% to GAS phages, with the vast majority of proteins highly homologous to phage Φ315.1 (4). Comparative genomics on 15 GAS phages has grouped them into six phage families (7). Given the high degree of genetic relationship and synteny in the molecular organization of Φ3396 and Φ315.1, Φ3396 is most likely a member of the 315.1-like phage family. Nonetheless, while sharing overall synteny, Φ3396 differs significantly from the 315.1 phage in the DNA replication module, host lysis module, and lysogenic conversion regions, implicating modular recombination as contributing to variations in phages within the same family. The host lysis and tail genes provide the phage with a level of host specificity, as they mediate the interaction between the phage and the cell wall of the bacterial host. Heterogeneity within these regions between Φ315.1 and Φ3396 may be a consequence of its specificity to GGS cell surface components.

In addition to sharing overall genomic identity to the 315.1 phage family, the flanking genes of the prophage 3396 in the NS3396 genome were the same as observed for Φ315.1 in the MGAS315 genome (4). This high degree of conservation further supports the suggestion of a distinct chromosomal integration site for the 315.1 phage family (7), which may be a reflection of the high degree of amino acid sequence identity

(98%) between the integrase proteins of Φ3396 and the 315.1 phage family.

Using multilocus sequence typing, Kalia et al. suggested that genetic movement for housekeeping genes may have occurred predominantly from GAS to GGS (22). These streptococci share a few genes encoding essential virulence factors (9, 20, 21, 29, 32, 37, 40) that exhibit mosaic structures consistent with ongoing interspecies LGTs. Although largely unidirectional, LGTs from GGS to GAS have also been documented (29). Directionality of phage movement between species depends on many factors including surface characteristics and restriction-modification systems of the participating bacteria. Interestingly, we found evidence of the 315.1 family of phages only within GGS and GAS screened from the area of GAS endemicity (NT, Australia). In contrast, none of the GAS and GGS strains from an area where they are not endemic were positive for this phage family. Furthermore, all NT GAS isolates that were PCR positive for the Φ3396-like phage share the same left attachment sites as suggested for the 315.1 phage family, while there was sequence variation within the lysogenic conversion regions which are flanked by a common right attachment site, suggesting the presence of 315.1-like phage variants within this population. Many GAS phage families including the 315.1-phage family exhibit high sequence diversity within the lysogenic conversion region which harbors virulence entities such as the superantigenic exotoxins (2, 7). Confirmation of mitomycin inducibility and prophage excision of 315.1-like phage within the NT GAS isolates further indicates that phages may be functionally viable within this population. Collectively, these findings suggest that the 315.1-like phage family appears to contribute to the population structure of circulating GAS and GGS strains within the area of GAS endemicity of the NT.

The observation that 315.1-like phages are more common in GAS strains than in GGS strains recovered from the same geographic locale may suggest that the progenitor Φ3396 may have arisen first in GAS before lateral movement to GGS. However, given the relatively small numbers of strains involved, this cannot be determined with certainty. Nonetheless, such lateral, ongoing acquisition of phages between a recognized pathogen and a largely commensal bacterium may not only have drastic effects on the overall population structure of the genus but also result in rapid changes to the pathogenic potential of GGS.

ACKNOWLEDGMENTS

We acknowledge Thanh Tran for technical assistance. We thank Joan Faoagali and her colleagues at the Department of Microbiology, Royal Brisbane Hospital, Brisbane, Australia, for isolates and relevant information. We also thank Bart Currie (Menzies School of Health Research, Darwin) for the NT isolates.

This work was carried out with the financial support of the Australian Government's Cooperative Research Centre for Vaccine Technology and the National Health and Medical Research Council (Australia).

REFERENCES

1. **Aziz, R. K., R. A. Edwards, W. W. Taylor, D. E. Low, A. McGeer, and M. Kotb.** 2005. Mosaic prophages with horizontally acquired genes account for the emergence and diversification of the globally disseminated M1T1 clone of *Streptococcus pyogenes*. J. Bacteriol. **187:**3311–3318.
2. **Banks, D. J., S. B. Beres, and J. M. Musser.** 2002. The fundamental contribution of phages to GAS evolution, genome diversification and strain emergence. Trends Microbiol. **10:**515–521.
3. **Banks, D. J., B. Lei, and J. M. Musser.** 2003. Prophage induction and expression of prophage-encoded virulence factors in group A *Streptococcus* serotype M3 strain MGAS315. Infect. Immun. **71:**7079–7086.
4. **Beres, S. B., G. L. Sylva, K. D. Barbian, B. Lei, J. S. Hoff, N. D. Mammarella, M. Y. Liu, J. C. Smoot, S. F. Porcella, L. D. Parkins, D. S. Campbell, T. M. Smith, J. K. McCormick, D. Y. Leung, P. M. Schlievert, and J. M. Musser.** 2002. Genome sequence of a serotype M3 strain of group A *Streptococcus*: phage-encoded toxins, the high-virulence phenotype, and clone emergence. Proc. Natl. Acad. Sci. USA **99:**10078–10083.
5. **Bonfield, J. K., K. Smith, and R. Staden.** 1995. A new DNA sequence assembly program. Nucleic Acids Res. **23:**4992–4999.
6. **Brussow, H., C. Canchaya, and W. D. Hardt.** 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol. Mol. Biol Rev. **68:**560–602.
7. **Canchaya, C., C. Proux, G. Fournous, A. Bruttin, and H. Brussow.** 2003. Prophage genomics. Microbiol. Mol. Biol. Rev. **67:**238–276.
8. **Carver, T. J., K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell, and J. Parkhill.** 2005. ACT: the Artemis comparison tool. Bioinformatics **21:**3422–3423.
9. **Cleary, P. P., J. Peterson, C. Chen, and C. Nelson.** 1991. Virulent human strains of group G streptococci express a C5a peptidase enzyme similar to that produced by group A streptococci. Infect. Immun. **59:**2305–2310.
10. **Cohen-Poradosu, R., J. Jaffe, D. Lavi, S. Grisariu-Greenzaid, R. Nir-Paz, L. Valinsky, M. Dan-Goor, C. Block, B. Beall, and A. E. Moses.** 2004. Group G streptococcal bacteremia in Jerusalem. Emerg. Infect. Dis. **10:**1455–1460.
11. **Colon, A. E., R. M. Cole, and C. G. Leonard.** 1971. Lysis and lysogenization of groups A, C, and G streptococci by a transducing bacteriophage induced from a group G *Streptococcus*. J. Virol. **8:**103–110.
12. **Cunningham, M. W.** 2000. Pathogenesis of group A streptococcal infections. Clin. Microbiol. Rev. **13:**470–511.
13. **Davies, M. R., T. N. Tran, D. J. McMillan, D. L. Gardiner, B. J. Currie, and K. S. Sriprakash.** 2005. Inter-species genetic movement may blur the epidemiology of streptococcal diseases in endemic regions. Microbes Infect. **7:**1128–1138.
14. **Efstratiou, A.** 1997. Pyogenic streptococci of Lancefield groups C and G as pathogens in man. Soc. Appl. Bacteriol. Symp. Ser. **26:**72S–79S.
15. **Ekelund, K., P. Skinhoj, J. Madsen, and H. B. Konradsen.** 2005. Invasive group A, B, C and G streptococcal infections in Denmark 1999–2002: epidemiological and clinical aspects. Clin. Microbiol. Infect. **11:**569–576.
16. **Facklam, R.** 2002. What happened to the streptococci: overview of taxonomic and nomenclature changes. Clin. Microbiol. Rev. **15:**613–630.
17. **Ferretti, J. J., W. M. McShan, D. Ajdic, D. J. Savic, G. Savic, K. Lyon, C. Primeaux, S. Sezate, A. N. Suvorov, S. Kenton, H. S. Lai, S. P. Lin, Y. Qian, H. G. Jia, F. Z. Najar, Q. Ren, H. Zhu, L. Song, J. White, X. Yuan, S. W. Clifton, B. A. Roe, and R. McLaughlin.** 2001. Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. Proc. Natl. Acad. Sci. USA **98:**4658–4663.
18. **Hartas, J., M. Hibble, and K. S. Sriprakash.** 1998. Simplification of a locus-specific DNA typing method (Vir typing) for *Streptococcus pyogenes*. J. Clin. Microbiol. **36:**1428–1429.
19. **Hindsholm, M., and H. C. Schonheyder.** 2002. Clinical presentation and outcome of bacteraemia caused by beta-haemolytic streptococci serogroup G. APMIS **110:**554–558.
20. **Igwe, E. I., P. L. Shewmaker, R. R. Facklam, M. M. Farley, C. van Beneden, and B. Beall.** 2003. Identification of superantigen genes *speM*, *ssa*, and *smeZ* in invasive strains of beta-hemolytic group C and G streptococci recovered from humans. FEMS Microbiol. Lett. **229:**259–264.
21. **Kalia, A., and D. E. Bessen.** 2003. Presence of streptococcal pyrogenic exotoxin A and C genes in human isolates of group G streptococci. FEMS Microbiol. Lett. **219:**291–295.
22. **Kalia, A., M. C. Enright, B. G. Spratt, and D. E. Bessen.** 2001. Directional gene movement from human-pathogenic to commensal-like streptococci. Infect. Immun. **69:**4858–4869.
23. **Laupland, K. B., T. Ross, D. L. Church, and D. B. Gregson.** 2006. Population-based surveillance of invasive pyogenic streptococcal infection in a large Canadian region. Clin. Microbiol. Infect. **12:**224–230.
24. **Malke, H., B. Roe, and J. J. Ferretti.** 1985. Nucleotide sequence of the streptokinase gene from *Streptococcus equisimilis* H46A. Gene **34:**357–362.
25. **Maxted, W. R., and E. V. Potter.** 1967. The presence of type 12 M-protein antigen in group G streptococci. J. Gen. Microbiol. **49:**119–125.
26. **Poyart, C., G. Quesne, S. Coulon, P. Berche, and P. Trieu-Cuot.** 1998. Identification of streptococci to species level by sequencing the gene encoding the manganese-dependent superoxide dismutase. J. Clin. Microbiol. **36:**41–47.
27. **Rogers, Y. C., A. C. Munk, L. J. Meincke, and C. S. Han.** 2005. Closing bacterial genomic sequence gaps with adaptor-PCR. BioTechniques **39:**31–32, 34.
28. **Rolston, K. V.** 1986. Group G streptococcal infections. Arch. Intern. Med. **146:**857–858.
29. **Sachse, S., P. Seidel, D. Gerlach, E. Gunther, J. Rodel, E. Straube, and K. H. Schmidt.** 2002. Superantigen-like gene(s) in human pathogenic *Streptococcus dysgalactiae* subsp. *equisimilis*: genomic localisation of the gene encoding

streptococcal pyrogenic exotoxin G (speG(dys)). FEMS Immunol. Med. Microbiol. **34:**159–167.

30. **Smoot, J. C., K. D. Barbian, J. J. Van Gompel, L. M. Smoot, M. S. Chaussee, G. L. Sylva, D. E. Sturdevant, S. M. Ricklefs, S. F. Porcella, L. D. Parkins, S. B. Beres, D. S. Campbell, T. M. Smith, Q. Zhang, V. Kapur, J. A. Daly, L. G. Veasy, and J. M. Musser.** 2002. Genome sequence and comparative microarray analysis of serotype M18 group A *Streptococcus* strains associated with acute rheumatic fever outbreaks. Proc. Natl. Acad. Sci. USA **99:**4668–4673.
31. **Sonnhammer, E. L., and R. Durbin.** 1995. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene **167:**GC1–GC10.
32. **Sriprakash, K. S., and J. Hartas.** 1996. Lateral genetic transfers between group A and G streptococci for M-like genes are ongoing. Microb. Pathog. **20:**275–285.
33. **Stuart, J. G., and J. J. Ferretti.** 1973. Transduction of rifampin resistance in group A streptococci. J. Bacteriol. **115:**709–710.
34. **Sylvetsky, N., D. Raveh, Y. Schlesinger, B. Rudensky, and A. M. Yinnon.** 2002. Bacteremia due to beta-hemolytic *Streptococcus* group G: increasing incidence and clinical characteristics of patients. Am. J. Med. **112:**622–626.
35. **Tapp, J., M. Thollesson, and B. Herrmann.** 2003. Phylogenetic relationships and genotyping of the genus *Streptococcus* by sequence determination of the RNase P RNA gene, *mpB*. Int. J. Syst. Evol. Microbiol. **53:**1861–1871.
36. **Tettelin, H., V. Masignani, M. J. Cieslewicz, J. A. Eisen, S. Peterson, M. R. Wessels, I. T. Paulsen, K. E. Nelson, I. Margarit, T. D. Read, L. C. Madoff, A. M. Wolf, M. J. Beanan, L. M. Brinkac, S. C. Daugherty, R. T. DeBoy, A. S. Durkin, J. F. Kolonay, R. Madupu, M. R. Lewis, D. Radune, N. B. Fedorova, D. Scanlan, H. Khouri, S. Mulligan, H. A. Carty, R. T. Cline, S. E. Van Aken, J. Gill, M. Scarselli, M. Mora, E. T. Iacobini, C. Brettoni, G. Galli, M. Mariani, F. Vegni, D. Maione, D. Rinaudo, R. Rappuoli, J. L. Telford, D. L. Kasper, G. Grandi, and C. M. Fraser.** 2002. Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*. Proc. Natl. Acad. Sci. USA **99:**12391–12396.
37. **Towers, R. J., D. Gal, D. McMillan, K. S. Sriprakash, B. J. Currie, M. J. Walker, G. S. Chhatwal, and P. K. Fagan.** 2004. Fibronectin-binding protein gene recombination and horizontal transfer between group A and G streptococci. J. Clin. Microbiol. **42:**5357–5361.
38. **Vandamme, P., B. Pot, E. Falsen, K. Kersters, and L. A. Devriese.** 1996. Taxonomic study of lancefield streptococcal groups C, G, and L (*Streptococcus dysgalactiae*) and proposal of *S. dysgalactiae* subsp. *equisimilis* subsp. nov. Int. J. Syst. Bacteriol. **46:**774–781.
39. **Van Domselaar, G. H., P. Stothard, S. Shrivastava, J. A. Cruz, A. Guo, X. Dong, P. Lu, D. Szafron, R. Greiner, and D. S. Wishart.** 2005. BASys: a web server for automated bacterial genome annotation. Nucleic Acids Res. **33:** W455—W459.
40. **Walter, F., M. Siegel, and H. Malke.** 1989. Nucleotide sequence of the streptokinase gene from a group-G *Streptococcus*. Nucleic Acids Res. **17:** 1262.