JOURNAL OF BACTERIOLOGY, June 2003, p. 3473–3475 Vol. 185, No. 11
0021-9193/03/$08.00+0 DOI: 10.1128/JB.185.11.3473–3475.2003
Copyright © 2003, American Society for Microbiology. All Rights Reserved.

Genomic Structure of the *Salmonella enterica* Serovar Typhimurium DT 64 Bacteriophage ST64T: Evidence for Modular Genetic Architecture

Princess T. Mmolawa,[1,2,3] Horst Schmieger,[4] Carly P. Tucker,[1,5] and Michael W. Heuzenroeder[1]*

Infectious Diseases Laboratories, Institute of Medical and Veterinary Science,[1] and School of Pharmaceutical, Molecular and Biomedical Sciences,[5] The University of South Australia, Adelaide, South Australia, 5000, and Discipline of Microbiology and Immunology, Department of Molecular Biosciences, University of Adelaide, Adelaide, South Australia, 5005,[2] Australia; Faculty of Agriculture, Science and Technology, Department of Animal Health, University of North West, Mmabatho, South Africa, 2735[3]; and Institute for Genetics and Microbiology, University of Munich, D-80638, Munich, Germany[4]

Received 23 October 2002/Accepted 20 March 2003

The complete sequence of the double-stranded DNA genome of a serotype-converting temperate bacteriophage, ST64T, was determined. The 40,679-bp genomic sequence of ST64T has an overall GC content of 47.5% and was reminiscent of a number of lambdoid phages, in particular, P22. Inferred proteins of ST64T which exhibited a high degree of sequence similarity to P22 proteins (>90%) included the functional serotype conversion cassette, integrase, excisionase, Abc1, Abc2, early antitermination (gp24), NinD, NinH, NinZ, packaging (gp3 and gp2), head (with the exception of gp26, gp7, gp20, and gp16), and tail proteins. The putative immunity genes were highly related to those of *Salmonella enterica* serotype Typhimurium phage L, whereas the lysis genes were almost identical to those of *S. enterica* serovar Typhimurium PS3.

Salmonella enterica serovars are known to harbor many temperate bacteriophages. Most of these belong to the P22 branch of the lambdoid family and are able to facilitate horizontal genetic transfer by transduction (11). Two temperate bacteriophages, ST64T and ST64B, were induced from *S. enterica* serovar Typhimurium DT 64 (8). ST64T could be propagated on various serovars and phage types. It could mediate generalized transduction. The genome of ST64T was completely sequenced in both directions and characterized. The genomic sequence was 40,679 bp in size with an overall GC content of 47.5%, similar to the GC content of P22 (47.1%), which is in contrast to the 52% GC content of its *S. enterica* host (9). The putative open reading frames (ORFs) were determined, and inferred protein sequences were scanned for similarity against the GenBank database by using the BlastP program (1, 2). Direction, order, and base sequences of many ORFs showed clearly that ST64T is a member of the lambdoid phage family.

Evidence of a modular genome in ST64T. The ST64T genome is a striking example of a modular genome composition. Figure 1 demonstrates that the ST64T genome is a genetic mosaic composed of gene modules, which is typical among members of this family. This allowed inferred functions and designations to be assigned to many genes based upon similarity with characterized genes from other phages, in particular, P22. In view of this, the P22 nomenclature was adopted for ST64T ORFs. The morphopoetic gene products (gp3 through gp9) and some other segments are almost identical with those of P22, whereas the three proteins involved in cell lysis and encoded by the lysis genes *13*, *19*, and *15* resemble those of phage PS3. The immunity C region is identical with that of phage L. The deduced product of gene *12* corresponds to that of the *Escherichia coli* lambdoid phage HK022, whereas gp18, which functionally cooperates with gp12, exhibits only 66% similarity with the corresponding HK022 product, which is the most similar protein in the database. The antiterminator gp23 shows a high level of similarity (81%) with the corresponding gene product of another lambdoid phage, *E. coli* phage 21. Indeed, ST64T is composed of genome segments typical of at least five different members of the lambdoid phage family.

While most genes or groups of genes (e.g., the packaging and morphopoetic genes) correspond as entities with a specific genomic source and, therefore, represent modules, it is interesting that for gp20, only the C-terminal region is identical with the corresponding protein of phage P22 (data not shown). The N-terminal region, which possibly encodes a separate domain, could not be aligned to a known sequence. Examination of the sites of sequence transitions or the putative sites of ancestral recombination indicated that the majority of transitions are located at gene boundaries. A comparison between ST64T and P22 at the end of gene *ninZ* and the start of gene *23* in both phages revealed the presence of such a transition. The putative *ninZ* gene of ST64T shares 93% identity in nucleotide sequence with *ninZ* of P22. However, *23*, which is the next gene for both ST64T and P22, has no sequence similarity (see Fig. 1). The similarity between the shared *ninZ* genes persists until the first base after the termination codon, and then there is complete diversion of sequence (data not shown). The putative gene *23* of ST64T appears to be closely related to gene *Q* of phage 21 as judged by sequence similarity (81% sequence identity). A similar scenario has been reported for ORF146 of λ and gene *61* of HK97 (5).

ST64T has an incomplete *immI* region. In contrast to phage lambda and most other known lambdoid phages, P22 carries,

* Corresponding author. Mailing address: P. O. Box 14, Rundle Mall, Adelaide, South Australia, 5000, Australia. Phone: 618 8222 3275. Fax: 618 8222 3543. E-mail: heuzenroeder@imvs.sa.gov.au.

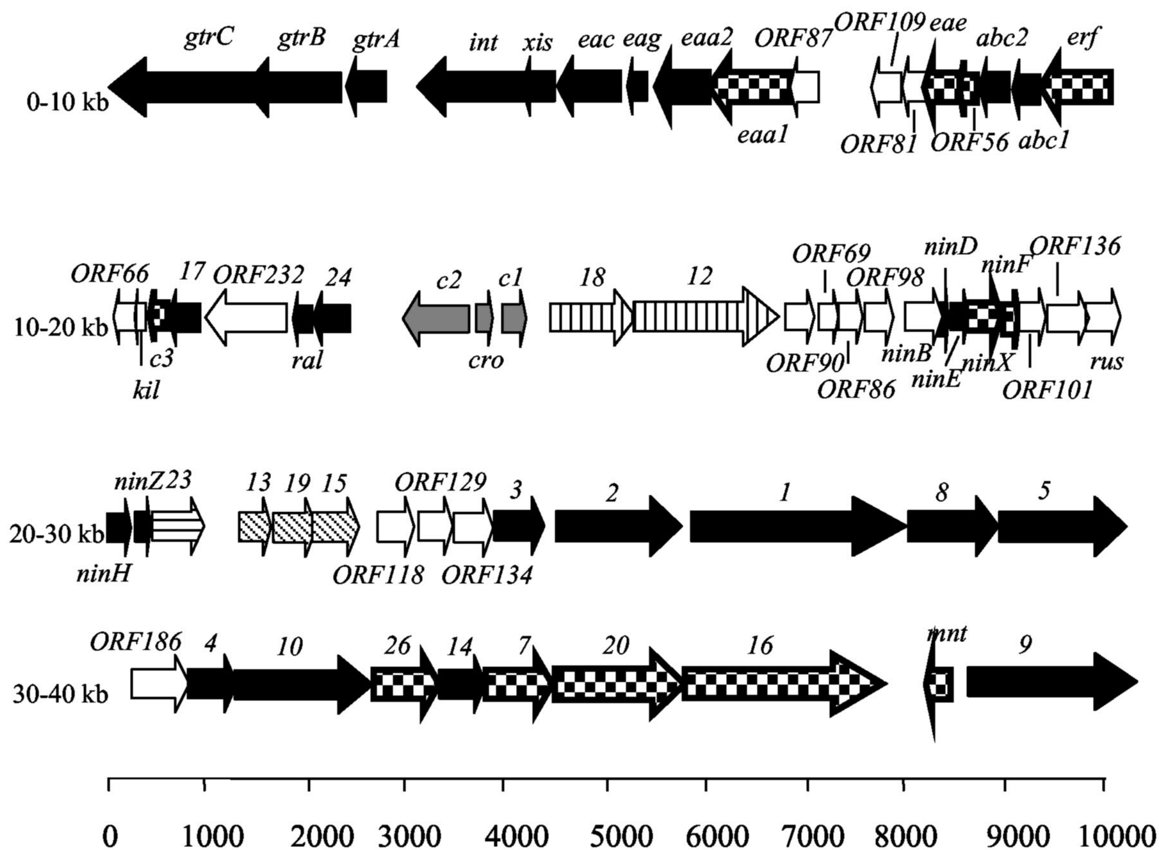

FIG. 1. Schematic representation of the phage ST64T genome indicating putative ORFs derived from the sequence. Similarities to other phage genes with known functions are indicated as shown below. The ORFs for which no genetic designation had been previously made are designated based upon the inferred size of the putative gene product: ■, >90% similarity to phage P22; ▩, <90% similarity to P22; ▤, similar to phage 21; □, no inferred function or similarity; ■, similar to phage L; ▥, similar to phage HK022; ▧, similar to phage PS3.

in addition to the *immC* region, a second region, *immI*, which expresses an antirepressor, Ant, and two repressors, Arc and Mnt, that regulate the expression of gene *ant*. Interestingly, phage ST64T has at the position of mnt_{P22} an ORF in the same orientation, with a deduced gene product that is only 60% similar to Mnt_{P22}. Mnt_{ST64T} has the same number of amino acids as Mnt_{P22} and also exhibits an almost identical hydropathy profile according to the Kyte-Doolittle algorithm (7) (Fig. 2). Therefore, it is highly likely that ST64T possesses an *mnt* allele. It could be speculated that the heteroimmune status of ST64T for P22 (8) might be a result of the expression of the mnt_{ST64T} allele. The ORFs corresponding to *arc* and *ant,* however, are not present. It is not known whether the mnt_{ST64T} allele is expressed. It is unknown by what mechanism this solitary gene of the *immI* system was introduced into the genome of ST64T. It may have been acquired separately from another prophage, or there was an ancestral phage that possessed the entire cassette, which has lost genes in the evolution of ST64T.

The serotype conversion cassette. Two PCR primers (forward, 5′ CCGCGTCATACCTGCGCTCACACGTCC 3′; reverse, 5′ GGGCCCGGTTCTGGCTCACCGTGG 3′) were designed to amplify the putative *gtrABC* seroconversion cassette and its putative upstream promoter using the GeneAmp XL PCR kit (Applied Biosystems). The 3,304-bp amplicon was ligated into the pCR2.1-TOPO vector and transformed into *E. coli* TOP10 chemically competent cells (Invitrogen). The cells were resuscitated in SOC broth and incubated for 1 h at 37°C. Aliquots were plated onto Columbia agar plates (special peptone, 2.3%; starch, 0.1%; sodium chloride, 0.5%; and agar, 1.1%) containing ampicillin (50 μg/ml), kanamycin (50 μg/ml), and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-

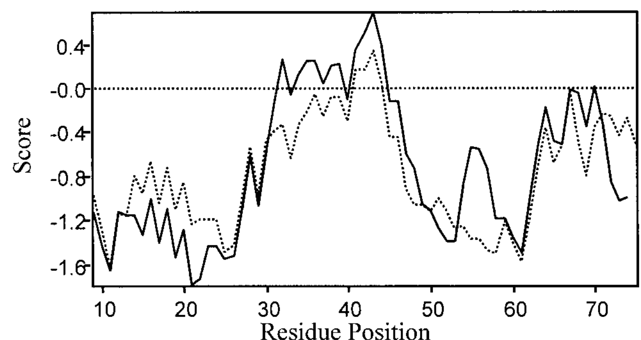

FIG. 2. Hydropathy prediction of the putative Mnt protein of ST64T compared to the Mnt protein of P22 using the Kyte-Doolittle hydropathy algorithm (7). The plot was produced using the OMIGA 1.1 program. The black line represents Mnt_{ST64T}, and the dotted line represents Mnt_{P22}.

Gal) (40 μg/ml). Plasmid DNA was isolated using the three-step alkaline lysis technique (10) and sequenced to determine the orientation of the insert relative to the vector. Plasmids representing the forward (pGTR-F) and reverse (pGTR-R) orientations of the insert were electroporated into serovar Typhimurium LT2 using standard procedures (3). Both serovar Typhimurium LT2 clones agglutinated with the anti-O1 serum, confirming that the serotype conversion cassette in ST64T is expressed from its own promoter and encodes the genes required for complete seroconversion.

The genomic similarity to other lambdoid phages. The ST64T genomic sequence data strongly suggest that serovar Typhimurium ST64T is a member of the P22-like lambdoid family (Fig. 1). One of the differences observed between ST64T and P22 is the genome size. The genome of ST64T is 40,679 bp, whereas that of P22 is 41,724 bp. This may be because ST64T lacks the other *immI* genes (*ant* and *arc*) as well as both superinfection exclusion genes, *sieB* and *sieA*. However, based upon restriction endonuclease digestion of the L genome, the ST64T genome size is much closer to the estimated size of L (40,650 ± 0.400 kb [4] and/or 40,500 kb [6]). The remarkably high sequence similarity in the *immC* region between these two phages (L and ST64T) suggests that they are related.

The genomic architecture of ST64T is similar to that of P22, and a number of regions are very similar to those of P22, including a region encompassing the O-antigen conversion genes. This similarity also extends to the genes involved in integration and excision that are transcribed in the opposite direction to the structural genes and have 98 to 100% sequence identities to P22. Other inferred gene products predicted to be involved in homologous recombination and packaging as well as many of the gene products involved in morphogenesis were >90% identical to similar proteins of P22. In contrast, the immunity and the lysis genes were similar to those of phages L and PS3, respectively. This is not totally surprising, since both L and PS3 are also P22-like phages carried by serovar Typhimurium. In addition, L has been shown to be a close relative of P22 (4). Nevertheless, the high level of sequence similarity of the putative replication genes and the late gene regulator *23* with analogous genes in the "true" lambdoid bacteriophages (HK022 and 21) demonstrates that ST64T, like most of the lambdoid phages, is a genetic mosaic made up of different modules, groups of genes which together perform a particular function.

Nucleotide sequence accession number. The complete nucleotide sequence of the ST64T bacteriophage has been deposited with GenBank and has been assigned accession number AY052766.

We thank Arthur Mangos and the Molecular Pathology Sequencing Laboratory for assisting with sequencing the ST64T genome and Andrew Kropinski for providing his Online Analysis Tools website, which was valuable during the analysis of sequence data. The assistance of the Australian Salmonella Reference Centre in phage typing and interpretation of results is gratefully acknowledged.

We thank the Australian Agency for International Development (AusAID) and the Rural Industries Research and Development Corporation (chicken meat program) for their generous financial support.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–4022.
3. **Dower, W. J., J. F. Miller, and C. W. Ragsdale.** 1988. High efficiency of transformation of Escherichia coli by high voltage electroporation. Nucleic Acids Res. **16:**6127–6145.
4. **Hayden, M., M. B. Adams, and S. Casjens.** 1985. Bacteriophage L: chromosome physical map and structural proteins. Virology **147:**431–440.
5. **Juhala, R. J., M. E. Ford, R. L. Duda, A. Youlton, G. F. Hatfull, and R. W. Hendrix.** 2000. Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. J. Mol. Biol. **299:**27–51.
6. **Karlovsky, P., J. Soska, and J. Reich.** 1984. Physical map of the bacteriophage L (*Salmonella typhimurium*). FEMS Microbiol. Lett. **25:**117–120.
7. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
8. **Mmolawa, P. T., R. Willmore, C. J. Thomas, and M. W. Heuzenroeder.** 2002. Temperate phages in *Salmonella enterica* serovar Typhimurium: implications for epidemiology. Int. J. Med. Microbiol. **291:**633–644.
9. **Ochman, H., and J. G. Lawrence.** 1996. Phylogenetics and the amelioration of bacterial genomes, p. 2627–2637. *In* F. C. Neidhardt, R. Curtis III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed., vol. 2. ASM Press, Washington, D.C.
10. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Laboratory Press, Cold Spring Harbor, N.Y.
11. **Schicklmaier, P., and H. Schmieger.** 1995. Frequency of generalized transducing phages in natural isolates of the *Salmonella typhimurium* complex. Appl. Environ. Microbiol. **61:**1637–1640.