# The Genome Sequence of *Yersinia pestis* Bacteriophage φA1122 Reveals an Intimate History with the Coliphage T3 and T7 Genomes

Emilio Garcia,[1] Jeffrey M. Elliott,[1] Erlan Ramanculov,[2] Patrick S. G. Chain,[1] May C. Chu,[2]* and Ian J. Molineux[3]

*Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, Livermore, California[1]; Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Fort Collins, Colorado[2]; and Molecular Genetics and Microbiology, Institute for Cell and Molecular Biology, University of Texas, Austin, Texas[3]*

The genome sequence of bacteriophage φA1122 has been determined. φA1122 grows on almost all isolates of *Yersinia pestis* and is used by the Centers for Disease Control and Prevention as a diagnostic agent for the causative agent of plague. φA1122 is very closely related to coliphage T7; the two genomes are colinear, and the genome-wide level of nucleotide identity is about 89%. However, a quarter of the φA1122 genome, one that includes about half of the morphogenetic and maturation functions, is significantly more closely related to coliphage T3 than to T7. It is proposed that the yersiniophage φA1122 recombined with a close relative of the *Y. enterocolitica* phage φYeO3-12 to yield progeny phages, one of which became the classic T3 coliphage of Demerec and Fano (M. Demerec and U. Fano, Genetics 30:119-136, 1945).

*Yersinia pestis* is the etiologic agent of plague, which is transmitted by a bite from infected fleas to their mammalian hosts. Humans become at risk for infection with *Y. pestis* when they are bitten by infective fleas or exposed through cuts in the skin while handling sick or dead animals or by inhaling aerosolized particles containing live organisms. Bubonic plague is the most frequently seen form in humans; it occurs when *Y. pestis* migrates from the fleabite site to the nearest lymph node and replicates there. The lymph node becomes inflamed and tender, and in this form it is called a bubo. From the bubo, *Y. pestis* can spill over into the bloodstream (septicemic plague) and, if untreated or unresolved, can enter the lungs, resulting in pneumonic plague. Although this rarely occurs, pneumonic plague is associated with high fatality as the lungs fill with fluid and the patient coughs bloody sputum and rapidly succumbs from an inability to breathe. Coughing also generates *Y. pestis* aerosols that rapidly spread and infect unprotected persons. These are the features that make plague so dangerous and historically feared and explain why plague has been used as a biowarfare agent (32).

**Y. pestis bacteriophages.** The period from the late 1920s through the early 1940s was an active time for investigators who recovered, identified, and characterized phages. Under laboratory conditions, *Y. pestis* has been shown to be a host for multiple phage strains from varied origins. D'Herelle described using lytic phage in 1919 as a treatment for plague (68). In 1927 Flu (22) recovered phages from canal waters in Leyden, Holland, that were active against *Y. pestis*, *Escherichia coli*, and *Shigella dysenteriae*. In 1929 Pokrovskaya recovered a lytic phage from infected (suslik) tissue that was specific to *Y. pestis* (54). Sugino in Japan and Advier in Senegal also reported *Y. pestis*-specific phages (2, 67). Girard confirmed the earlier data and suggested that phages could be developed as useful tools for the diagnosis and treatment of plague (23).

The most detailed studies of *Y. pestis* phages have been made by investigators in the former Soviet Union. The d'Herelle phage and many other strains isolated in the former USSR comprise four serovars. Serovar 1 is by far the most common and is represented by the d'Herelle and Pokrovskaya lytic phages; both of these phages exhibit T7-like growth curves and morphology (4, 74) and are probably members of the T7 phage family. Serovar 2 phages are P2-like (E. Garcia et al., unpublished data); they were originally defined by phages H (46) and L-413 C (34). (This phage H is not identical to the phage H described by Molnar and Lawton [41], which is discussed below.) Serovar 3 is represented by lysogenic phage Π (47), and serovar 4 is represented by the lysogenic phages Tal and 513 (48). The three lysogenic phage serovars are said to be heteroimmune.

Although early work of d'Herelle was directed towards curing bubonic plague by administration of phage (phage therapy), which was, by his own account, a success, it did not become the focus of subsequent studies (68). Rather, most attention concentrated on the isolation of phages that could specifically type all natural isolates of *Y. pestis* and on the adaptation of phages to grow more efficiently on other enteric bacteria (see, e.g., reference 35). However, it is often difficult to determine from the old literature exactly which phage was being used; "the bacteriophage" is not an uncommon description. In addition, the specific protocols for phage propagation and its use by different workers are not always clearly described. In some cases descriptions of phage growth are accompanied by efficiencies of plating and plaque morphologies determined by

using soft agar overlays, whereas in others phage growth is scored by the routine test dilution method (1), as plaques on the surface of agar plates, or even by simple spot tests (which measure only cell killing and not phage growth).

**History of bacteriophage φA1122.** In 1933 Advier (2) isolated a *Y. pestis* phage from blood of a patient with a clinical case of bubonic plague. The phage was subsequently adapted to the avirulent *Y. pestis* strain A1122 (33). A stock of this phage was obtained in 1945 by Lazarus and Gunnison from the Pasteur Institute of Dakar, Senegal. This "parent phage" was passaged on A1122 in rich media at 37°C, using both liquid and semisolid media, and plaque purified several times and was then described as "purified phage" (35). Gunnison et al. (25) appear to have renamed the purified phage as "P phage" when they reported that it could be used to differentiate *Y. pestis* and *Yersinia pseudotuberculosis*. At 37°C P phage grows on both species, but at 20°C it only grows on *Y. pestis*. In practice, differentiation between *Y. pestis* and *Y. pseudotuberculosis* by phage growth is robust up to 28°C.

A paper strip test containing phage was subsequently developed by Cavanaugh and Quan as a test for the rapid identification of *Y. pestis* (10). Curiously, given the importance of that communication, the authors referred to this diagnostic agent only as "the bacteriophage." The uncertain provenance of the Cavanaugh and Quan phage, which was said to be isolated from sewage at the Hooper Foundation, University of California, Berkeley, was presented by Molnar and Lawton (41). Those authors arbitrarily named the Cavanaugh and Quan phage "phage H." Although phage H had properties similar to those of phage P (25), which was being used at the Hooper Foundation at the same time that phage H was isolated, there was no proof that the two phages were identical. Like phage P, phage H is used at 20°C for diagnosing *Y. pestis* (10, 26) and was used by the U.S. Army Medical Research Institute of Infectious Diseases as their standard phage for identifying *Y. pestis* (see, e.g., reference 26). In recent years, the U.S. Army Medical Research Institute of Infectious Diseases has been using the φA1122 stock of the Centers for Disease Control and Prevention (CDC) for lysis of *Y. pestis* cultures (Pat Worsham, personal communication). The stock of phage whose genome sequence is presented here also originated from the Hooper Foundation. It was transferred to the Vector-Borne Diseases Division of the CDC in Fort Collins, Colo., in 1968. The documentation that accompanied this stock strongly implies, although it does not explicitly state, that it was the Gunnison P phage. Because of its uncertain history, we have renamed the phage φA1122, reflecting the host *Y. pestis* strain used for routine propagation. Strain A1122 is used by the CDC as the reference *Y. pestis* strain (12). Lysates derived from the original Hooper Foundation stock have been in continual use and widely distributed since 1968 by the CDC as one of the diagnostic tests for the identification of *Y. pestis*. φA1122 grows on all but two natural isolates of *Y. pestis* of the thousands tested by the CDC (M. C. Chu, unpublished data). The former Soviet Union (now Russian Federation and Central Asian Republics) antiplague networks similarly use the T7-related Pokrovskaya phage and the P2-related L-413 C phage for diagnostic purposes.

The φA1122 genome sequence determined as part of this study shows that φA1122 is very closely related to T7, although a significant portion of the genome is almost identical to the equivalent part of coliphage T3. Analysis of the sequence data suggests that T3, one of the classic set of T phages described by Demerec and Fano (14), was derived from a recombination event involving two different yersiniophages.

## MATERIALS AND METHODS

**Bacteriophages, bacteria, and plasmids.** φA1122 is the standard phage used by the CDC in identifying wild or clinical isolates of *Y. pestis*. T7 and T3 are the laboratory stocks of I. J. Molineux. *Y. pestis* A1122, the CDC laboratory standard avirulent, *Y. pestis* bv. *orientalis* strain lacking the 70-kb pCad virulence plasmid, was used for φA1122 propagation. Host range mutants of φA1122 were selected on strain IJ511 (*E. coli* K-12 Δ*lacX74 supE44 galK2 galT22 mcrA rfbD1 mcrB1 hsdS3*) or BL21 (*E. coli* B *hsdS* Gal⁻). DH5α and DH5α/F' *lacI*�q, used for phage library propagation, were purchased from (Gibco-BRL Life Technologies). Other bacterial strains used were from the collection of I. J. Molineux and are described in Table 2. The plasmid pAR3685 (a gift from A. H. Rosenberg and F. W. Studier, Brookhaven National Laboratories) contains T7 φ17 and gene *17* cloned at the *Bam*HI site of pBR322 in the opposite orientation to P*tet*.

Cells were grown in brain heart infusion (Difco) or Luria-Bertani medium, containing 100 μg of ampicillin per ml when appropriate. Plasmid transformation was by electroporation. Unless otherwise specified, cultures of A1122 were grown at 30°C.

**Phage DNA library and DNA sequencing.** For genome sequence determination, φA1122 was grown on A1122 by using a single plaque derived from the original stock provided to the CDC. A large-scale lysate was clarified by centrifugation and then concentrated by precipitation with 7.5% polyethylene glycol 8000 (Sigma). The pellet was resuspended and purified by CsCl density gradient centrifugation. DNA was extracted with phenol-chloroform (24:1, vol/vol) and precipitated with ethanol.

Phage DNA was subjected to hydroshearing (Gene Machines; Genomics Instrumentation Services, Inc., San Carlos, Calif.) to yield 0.5- to 1 kb fragments, which were cloned into dephosphorylated, *Sma*I-digested pUC18 and M13mp18 vectors. A 5-times-redundant M13 library (four 96-well plates) and a ~12-times-redundant pUC-based library (five 96-well plates) were constructed. The majority of sequencing used Big Dye chemistry (PE Applied Biosystems), ABI PRISM 373 or 377 sequencers, and primers −21m13 and m13rp1 (Operon Technologies, Inc.). Approximately 1,300 reads were sequenced and assembled, resulting in a final coverage of about 16×. Base calling and assembly of sequences used Phred/Phrap (20, 21) and Consed (24). Ambiguities were resolved and the genome ends were determined directly from phage genomic DNA by cycle sequencing with appropriately designed primers.

**Sequence annotation.** The final sequence was searched against the current protein and nucleotide databases (http://www.ncbi.nlm.nih.gov/) by using BLAST (3). Alignment of open reading frames to T7, T3, and φYeO3-12 sequences was done with Lasergene (DNAstar, Madison, Wis.), GeneWorks 2.45 (IntelliGenetics Inc., Mountain View, Calif.), the Artemis Comparison Tool (http://www.sanger.ac.uk/Software/Artemis/), and Alion:pairwise alignment (http://motif.stanford.edu/alion/). Computer-based predictions were then evaluated manually.

**Nucleotide sequence accession number.** Nucleotide sequence data for φA1122 have been deposited in GenBank under accession no. AY247822.

## RESULTS AND DISCUSSION

**Genome organization.** The φA1122 genome consists of 37,555 bp of DNA that includes direct terminal repeats of 148 bp. A cursory inspection of the sequence reveals that the phage belongs to the T7 family and that the genome is colinear with those of T7, T3, and the *Yersinia enterocolitica* phage φYeO3-12. The G+C content of the φA1122 genome is 48.3%, and that of its host *Y. pestis* is 47.6% (15, 51). For comparison, the G+C contents of the coliphages T7 and T3 are 48.4 and 49.9%, respectively, while that of *E. coli* K-12 is 50.8%. In total, 51 gene products are predicted in the φA1122 genome. The genome contains 46 distinct open reading frames plus three predicted overlapping genes in a second reading frame (Table 1; Fig. 1). In addition, and as in the T7 genome sequence, an
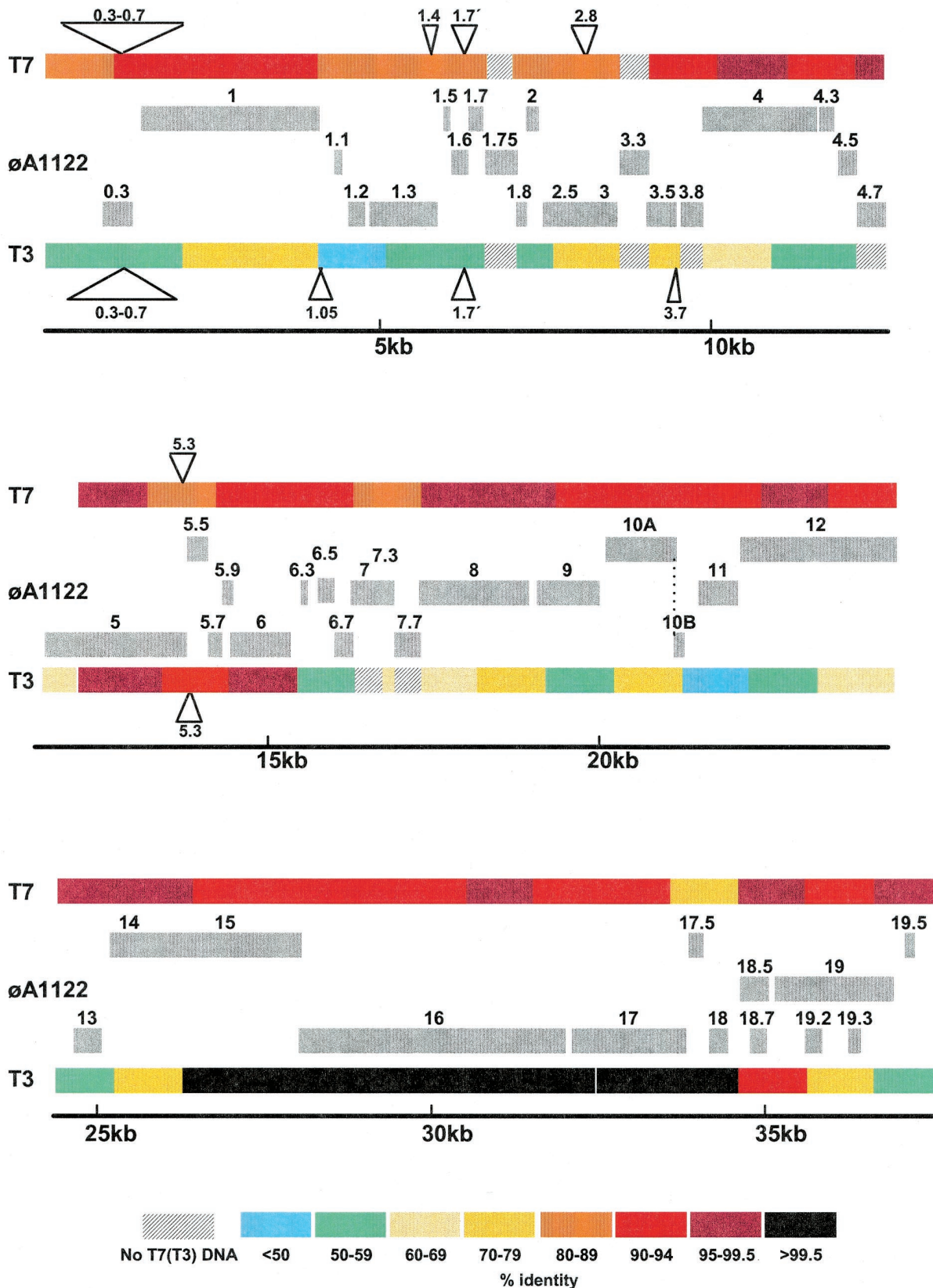
FIG. 1. Genetic organization of the φA1122 genome and comparison with T7 and T3. T7 and T3 genes that are not present in φA1122 are shown as triangles, and φA1122 genes that are not present in T7 or T3 are shown by gray boxes. Percent similarities of the φA1122 genome with those of T7 and T3 were calculated by using 1-kb segments, after removing the regions indicated where the phages under consideration contain significant unequal genetic content.

internal initiation site provides the C-terminal, helicase-only (gp4B) portion of the φA1122 gene *4* helicase-primase, and a translational −1 frameshift near the end of the coding sequence for the major capsid protein gp10A allows synthesis of the minor form gp10B. The known regulatory nucleotide sequences allowing synthesis of T7 gp4B and gp10B are totally conserved with those in the φA1122 genome, and it can thus be assumed that the latter phage also makes two products from both genes *4* and *10*. In contrast, a homolog to the putative T7 gene *4.1*, which overlaps gene *4*, is unlikely to be expressed in φA1122, as the open reading frame lacks a convincing ribosome-binding site. Neither T3 nor φYeO3-12 is thought to code for gene *4.1*, and it may not be a real gene in the T7 phage family. φA1122 also lacks both an appropriate Shine-Dalgarno sequence and a start codon for a homolog to T7 gene *4.2*, another putative gene that overlaps gene *4*. It is unlikely that φA1122 gene *4.2* is expressed. However, homologs of T7 gene *4.2* may be expressed in T3 and in φYeO3-12.

The φA1122 genome is ~2.5 kb shorter than that of T7; the major differences in genetic content are due to a 1.75-kb deletion that fuses the counterparts of T7 genes *0.3* and *0.7* and to the lack of genes *1.4*, *2.8*, and *5.3* and the first 120 codons of gene *1.7*. Partially countering these deletions (still relative to T7) are two genes, *1.75* and *3.3*, that have no positional T7 counterparts. Small deletions also affect the φA1122 terminal repeats and genes *1.3* and *7.3*. However, genes *1.3* and *7.3* should both retain activity; the host range mutants of φA1122 described below plate on *E. coli* strains deficient in DNA ligase, which renders the phage ligase essential (Table 2). Furthermore, the T7 gene *7.3* protein is an essential component of the phage particle (P. Kemp and I. J. Molineux, unpublished data) and likely has a similar role in φA1122. Table 1 summarizes the regulatory and coding elements of the φA1122 genome. Aside from the putative φA1122 gene products that have no counterparts in the T7 genome, the predicted sequences of most other proteins exhibit >80% identity with the homologous T7 protein. Three, gp1.5, gp3.8, and gp4.3 (all of unknown function), are 100% conserved. In contrast, gp5.5 retains 46% identity, and gp1.7, which is present only in a truncated form, retains only 44%. Although the similarity to most T3 proteins is generally much lower than that to T7 proteins, four φA1122 proteins (gp5.7, gp17.5, gp18, and gp18.5) are 100% identical in sequence to their T3 counterparts.

**Regulatory elements.** The φA1122 genome contains the same three major early promoters for *E. coli* RNA polymerase (RNAP) as T7 (Table 1); the minor C promoter and the left-ward-facing A0 (or D) promoter are also present. The minor B promoter of T7 is missing as a consequence of the fusion of φA1122 genes *0.3* and *0.7*. The 17 promoters for φA1122 RNAP are found in positions equivalent to those in T7. Several are identical in sequence to their positional T7 counterparts, but the two promoters suggested to have a role in DNA replication, φOL and φOR, are predicted to have significantly different activities. Unlike that of T7, φA1122 φOL contains the consensus nucleotide −11G in the polymerase recognition domain and should thus be a stronger promoter (31) than T7 φOL. In contrast, φA1122 φOR is not the typical strong promoter found in the class III region of the T7 genome. φA1122 φOR contains a −5G change from the consensus, one that is

also found in the weak class II promoter φ*1.3*. Whether φA1122 φOR is also weak in vivo has not been determined experimentally, but in vitro the −5G change from the consensus reduces T7 promoter activity threefold (31).

The early transcriptional terminator TE, which defines the end of the region transcribed by the host RNAP, and the two phage polymerase terminators Tφ and CJ are all highly conserved in the T7 and φA1122 genomes. All are likely to function in vivo at an efficiency in φA1122 comparable to that during a T7 infection. Most of the RNase III recognition sites found in the T7 genome are also highly conserved in φA1122. No counterpart to T7 R*0.5* is present because of the deletion that fuses φA1122 genes *0.3* and *0.7* but R*0.3* and R*1.1* are identical in the two genomes, and R*18.5* is identical to T3 R*18.5*. In addition, relative to their T7 counterparts, *R1* and *R3.8* have only a single change in the top loop of the predicted imperfect hairpin that characterizes an RNase III recognition site. There are also minor differences between the R*4.7* elements of the two genomes; only R*1.3*, R*6.5*, and R*13* warrant additional mention.

Alone among the RNase III recognition structures, T7 R*1.3* contains two sites, separated by 29 bases, for cleavage by the enzyme (56). Both sites are used in vivo; the downstream site produces a longer gene *1.1-1.2* RNA whose 3′ end can base pair to the ribosome-binding region of gene *1.1* and thereby retroregulate translation (59). Although φA1122 R*1.3* has a significantly shorter bottom duplex stem than T7 R*1.3*, the predicted overall structure of the enzyme recognition site suggests that both cleavage sites are retained (Fig. 2). Furthermore, retroregulation of gene *1.1* expression is likely conserved, as the 3′ end of the longer *1.1* mRNA retains base-pairing properties with the *1.1* ribosome-binding site.

φA1122 R*6.5* differs from its T7 counterpart at two positions; it lacks a base in the 5′ part of the upper duplex stem, creating an additional bulge loop, and it also contains the more canonical sequence CUUUAU|G for the predicted site of cleavage by RNase III. The predicted structure of φA1122 R*13* differs from that of T7 in that it can be drawn as a perfect stem-loop (albeit with a large loop) rather than the bulged structures common to all T7 and other predicted φA1122 RNase III sites. RNase III cleaves T7 R*6.5* efficiently but is much less effective at cleaving R*13* (18), but any effects on cleavage by RNase III caused by the sequence differences between R*6.5* and R*13* of T7 and φA1122 cannot be predicted without experimental data. However, it should be noted that T7 grows essentially normally in *E. coli* strains that lack RNase III (17) and that *E. coli* RNase III is 92% identical to the *Y. pestis* enzyme. φA1122 host range mutants plate with normal efficiency in *E. coli rnc* mutants (Molineux, unpublished observations), and thus altered cleavage patterns between T7 and φA1122 at R*6.5* and R*13* are unlikely to be highly significant in vivo.

Genetic elements known or suggested to be involved in DNA replication include the terminal repeats TR, which, aside from a 13-bp deletion, are 95% identical to T7; the hairpin structure proposed to enable duplication of TR during packaging (13) is completely conserved. The primary origin of replication *ori* is identical in sequence to T7 *ori*, and both SRL and SRR, the arrays of imperfect repeats adjacent to each TR that

TABLE 1. Genomic comparison of φA1122 with T7 and T3[a]

| Nucleotide positions | Regulatory element | Gene | Translation initiation region[b] | No. of amino acids | T7 protein similarity (%) | T3 protein similarity (%) | T7 homolog function | Comments |
|---|---|---|---|---|---|---|---|---|
| 1–148 | TR | | | | | | Concatemer formation | Direct terminal repeat; 13-bp deletion, otherwise 95% identical to T7 |
| 149–155 | CJ | | | | | | DNA packaging (73) | Lysozyme-dependent phage RNAP terminator (38); position 162/163 likely 3′ end of RNA |
| 247–212 | A0 | | | | | | Phage promoter | Leftward E. coli promoter |
| 357–379 | φOL | | | | | | Major early E. coli promoter | Change from consensus: +2A |
| 465–501 | A1 | | | | | | Major early E. coli promoter | |
| 590–627 | A2 | | | | | | Major early E. coli promoter | 19-bp insertion in T7 from −4 to +15 |
| 695–731 | A3 | | | | | | RNase III site | |
| 834–882 | R0.3 | | | | | | | Identical to T7 R0.3 |
| 906–1328 | | 0.3–0.7 fusion | GCACGAGGTAACACAAGATGGCT | 140 | 96, 89 | Insignificant | Inactivates type I restriction; no protein kinase activity | First 108 aa[c] of protein are 96% identical to T7 gp0.3; last 35 aa are 89% identical to T7 gp 0.7; homologs of T7 0.4, R0.5, B promoter, 0.5, and 0.64/B are not present |
| 1306–1341 | C | | | | | | Minor E. coli promoter | |
| 1330–1376 | RJ | | | | | | RNase III site | |
| 1399–4050 | 1 | 1 | ACCGGAAGAGAGGCACTAAATGAAC | 883 | 98 | 81 | RNAP | T7 promoter specificity |
| 4059–4081 | φ1.1A | | | | | | Phage promoter | Changes from consensus: +2A + 4G |
| 4075–4128 | R1.1 | | | | | | RNase III site | Identical to T7 R1.1 |
| 4133–4155 | φ1.1B | | | | | | Phage promoter | Identical to T7 φ1.1B; changes from consensus: +2A, +3G, +4A |
| 4149–4275 | ori | | | | | | Primary origin of replication | Identical to T7 ori (60) |
| 4235–4366 | | 1.1 | AAGAGAGGACTTTAAGTATGCGT | 43 | 79 | 31 | Unknown | Conserved in T7, T3, φYeO3-12: like T3, translational coupling to 1.2 expression less likely than for T7. |
| 4368–4625 | | 1.2 | CTGGGAGGGTCAATAAAGATGGGT | 85 | 88 | 44 | Host dGTPase inhibitor; F exclusion target | Contains Q28; T7 Y28N nontoxic |
| 4623–4645 | φ1.3 | | | | | | Phage promoter | Identical to T7 φ1.3; changes from consensus: −5G, +5C |
| 4635–4691 | R1.3 | | | | | | RNase III site | The two cleavage sites for RNase III (56) and retroregulation of 1.1 expression (58) likely conserved with T7 |
| 4707–5729 | | 1.3 | ACATGGAGATAAACATTATGATG | 340 | 85 | 71 | DNA ligase | In-frame deletion, also found in T3 and φYeO3-12 gp1.3, corresponding to T7 aa 298–316 |
| 5737–5763 | TE | | | | | | Early transcription terminator | |
| 5767–5789 | φ1.5 | | | | | | Phage promoter | Identical to T7 φ1.5; changes from consensus: −2A, +3A, +4G, +6T; as in T3 and φYeO3-12, no T7 1.4 homolog |
| 5797–5886 | | 1.5 | TAAAGGAGGTACACATCATGTAC | 29 | 100 | 52 | Unknown | Also present in φYeO3-12 |
| 5884–5906 | φ1.6 | | | | | | Phage promoter | Identical to T7 φ1.6; changes from consensus: −2A, +3A, +4G, +5A, +6C |
| 5912–6172 | | 1.6 | ACTAAAGGAGACACTATATGTTT | 86 | 95 | 52 | Unknown | Also found in φYeO3-12 |
| 6172–6402 | | 1.7 | ATCAAGGAGGYGTTCTGATGGGC | 76 | 44 | 43 | Unknown | Similarities confined to C-terminal 75 aa of T7 gp1.7 (196 aa) and 65 aa of T3 gp1.7 (163 aa); 54% similar to C terminus of Listeria phage PSA gp52. |

| Position | Site/promoter | Gene no. | Sequence | aa | % | % | Function | Comments |
|---|---|---|---|---|---|---|---|---|
| 6383–6853 | | 1.75 | GCGTCTGGAGGAAGCAAATGAAA | 156 | | | | Possible downstream GTG start (position 6449); HNH endonuclease family protein; similarity to *S. agalactiae* serotype V prophage Sa2 |
| 6804–6956 | | 1.8 | TCATAGGATAAATCATATGAGA | 50 | Insignificant | 88 | Unknown | |
| 6958–7152 | φ2.5 | 2 | ATCGAGAGGTTACTGATATGTCA | 64 | 44 | 95 | Inhibits *E. coli* RNA | As in T3 and φYeO3-12, no positional homolog of T7 2.8 |
| 7150–7172 | | | | | | | Phage promoter | Identical to T7 φ2.5; changes from consensus: −1T, +1A, +4G, +5A |
| 7218–7916 | | 2.5 | TAAAGGAGATTAATATTATGGCT | 232 | 80 | 97 | SSB | 88% similar to φYeO3-12 gp3 |
| 7917–8372 | | 3 | CAGACGGAGACTTCTCAAGTGGCT | 151 | 82 | 85 | Endonuclease I | |
| 8369–8827 | | 3.3 | GCGAAAGGAGGTAAGAAGTGATA | 152 | | | | 40% similar to φA1122 gp7.7; 43% similar to 17 gp2.8; 31% similar to T7 and φYeO3-12 gp5.3; HNH endonuclease family protein |
| 8799–9254 | | 3.5 | CAAGTGGAGGAATTTAAAATGGCA | 151 | 90 | 93 | Lysozyme | N-acetylmuramyl amidase; inhibits phage RNAP |
| 9256–9288 | φ3.8 | 3.8 | | | | | Phage promoter | Identical to T7 φ3.8; changes from consensus: −13T, −12G, 11A, −2A |
| 9257–9305 | R3.8 | | | | | | RNase III site | |
| 9318–9683 | | | GCATTGGAGATCAAATAATGCGC | 121 | | 100 | Unknown | Similarity to T3 gp5.3 HNH endonuclease family protein. |
| 9658–11358 | 4A/4B | 4A/4B | AACTAGGAGGGAATTGCATGGAC/CCTCAGGAGGTAAACCAATGACG | 566/503 | 82 | 99 | Primase/helicase | As in T7, internal in-frame start yields gp4B; homologs to putative overlapping T7 genes 4.1 and 4.2, T3 gene 4.1, φYeO3-12 genes 4.1 and 4.15, lack appropriate Shine-Dalgarno sequence and/or start codon. |
| 10747–10769 | φ4c | | | | | | Phage promoter | Identical to T7 φ4c; changes from consensus: −17C, −13C, 2A, +2A |
| 11377–11399 | φ4.3 | | | | | | Phage promoter | Change from consensus: −17C, −2A, +3A, +4G, +5A, +6C |
| 11404–11616 | | 4.3 | CACTAAAGGAGGAGACACCCATGTTC | 70 | 42 | 100 | Unknown | |
| 11636–11905 | | 4.5 | AAACAGGAGAAACCATTATGTCT | 89 | 73 | 98 | Unknown | |
| 11906–11954 | R4.7 | | | | | | RNase III site | |
| 11950–11972 | φ4.7 | | | | | | Phage promoter | Identical to T7 φ4.7; changes from consensus: −17C, −16T, −13T, +3A, +4G, +5A, +6T |
| 11979–12386 | | 4.7 | CTATAGGAGATATTACCATGCGT | 135 | 96 | 96 | Unknown | Not present in T3 or φYeO3-12 |
| 12405–14519 | | 5 | CAATAGGAGAAATCAATATGATC | 704 | 98 | 98 | DNA polymerase | Homologs to putative T3 and φYeO3-12 internal in-frame gp5B unlikely; putative overlapping homolog (13048–13137) to T3 gp5.1 possible; no positional homolog of gene 5.3 |
| 14539–14835 | | 5.5 | TCTATAGGAGAAATTATTATGAC | 98 | 90 | 46 | Growth on λ *rex*+ lysogens; binds *E. coli* H-NS | Like T3 and φYeO3-12, no obvious frameshift site for homolog of T7 5.5–5.7 fusion |
| 14835–15044 | | 5.7 | GTACGGGAGGTATTCTGATGTCT | 69 | 100 | 98 | Unknown | |
| 15044–15202 | | 5.9 | ATGGGAGGTTGTGTCTAagTCT | 52 | 98 | 96 | Inhibits RecBCD nuclease | |
| 15189–16091 | | 6 | GCTAGGAGGAGAAACTTAATGGCG | 300 | 92 | 97 | Exonuclease | |
| 16079–16162 | | 6.3 | GACAAGGAGATTTACCTGTGGAG | 27 | 38 | 85 | Unknown | |
| 16187–16258 | R6.5 | | | | | | RNase III site | T7 and T3 proteins are respectively, 38 and 37 aa |
| 16213–16235 | φ6.5 | | | | | | Phage promoter | |
| 16289–16543 | | 6.5 | TAAGAGGAATCTTTATCATGTTA | 84 | 55 | 94 | Unknown | |
| 16548–16814 | | 6.7 | TGGGGAGGATTGACCCTTATGTGT | 88 | 70 | 96 | Internal head protein (Kemp and Molineux, unpublished data) | Consensus |

TABLE 1—*Continued*

| Nucleotide positions | Regulatory element | Gene | Translation initiation region[b] | No. of amino acids | T7 protein similarity (%) | T3 protein similarity (%) | T7 homolog function | Comments |
|---|---|---|---|---|---|---|---|---|
| 16814–17215 | | 7 | TTTGGAGATAAGAAGTGATGTCT | 133 | 99 | | Host range | Not present in T3 or φYeO3-12 |
| 17219–17458 | | 7.3 | TTAAGGAGGTATAAGTTATGGGT | 79 | 76 | 55 | Tail protein (Kemp and Molineux, unpublished data) | Relative to T7 gp7.3, has N-terminal 11- and 9-aa internal deletions and C-terminal 49 aa are 100% identical; relative to T3 and φYeO3-12 gp7.3, has 25-aa N-terminal deletion |
| 17472–17864 | | 7.7 | TTAATCAGGAGGTTATCGTGGAA | 130 | 99 | | | HNH endonuclease family protein; not present in T3 or φYeO3-12 |
| 17864–19474 | | 8 | ACATGGAGACACATTTAATGGCT | 536 | 99 | 85 | Head-tail connector | Consensus |
| 19472–19494 | φ9 | | | | | | Phage promoter | Consensus |
| 19574–20488 | | 9 | TTAAGGAGACAATAATAATGGCT | 304 | 95 | 60 | Scaffolding protein | Consensus |
| 20502–20524 | φ10 | | | | | | Phage promoter | Consensus |
| 20587–21621 | | 10 | AAGAAGGAGACATACATATGGCT | 344/39 | 94/94 | 79/71 | Major and minor capsid proteins gp10A and gp10B | −1 frameshifting site for gp10B synthesis conserved with T7 |
| 20587–21779 | | Tφ | | 8 | | | Late transcription terminator | |
| 21787–21827 | Tφ | | | | | | | Identical to T7 Tφ |
| 21848–22438 | | 11 | AAAGGAGGAGGAACTATATGCGC | 196 | 97 | 80 | Tail protein | |
| 22459–24843 | | 12 | TCAATAAGGAGGCTCTAAATGGCA | 794 | 96 | 67 | Tail protein | |
| 24849–24899 | R13 | | | | | | RNase III site | |
| 24874–24896 | φ13 | | | | | | Phage promoter | Consensus |
| 24924–25340 | | 13 | TACGGGATGGTTTCTTATGATG | 138 | 98 | 50 | Head assembly protein | Poor Shine-Dalgarno sequence |
| 25345–25935 | | 14 | GAAAGGAGGATAAACCATATGTGT | 196 | 94 | 70 | Internal virion protein | |
| 25942–28185 | | 15 | CGGGGAGGTAATGAACTATGAGC | 747 | 98 | 97 | Internal virion protein | |
| 28212–32168 | | 16 | ACATAAGGAGGCGCCCTAAAATGGAT | 1318 | 97 | 99 | Internal virion protein | |
| 32166–32188 | φ17 | | | | | | Phage promoter | Consensus |
| 32241–33917 | | 17 | TACTTTAAGGAGGTCAAATGGCT | 558 | 86 | 98 | Tail fiber | |
| 33976–34179 | | 17.5 | CACACTAAGGAGGACACATGTTG | 67 | 94 | 100 | Holin | λ Rz and P22 homologs required for lysis in presence of divalent cations |
| 34184–34453 | | 18 | AACTAAGGGAGTAACTCTATGGAA | 89 | 92 | 100 | Small terminase subunit | |
| 34547–34990 | R18.5 | | | | | | RNase III site | Identical to T3 R18.5 |
| 34547–34990 | | 18.5 | AGTTAACGGGAGCCATTATGCTA | 147 | 91 | 100 | λ Rz homolog; complements P22 15 mutants (9) | Overlaps gene 18.5; λ Rz and Rz1 both required for lysis (70) |
| 34662–34913 | | 18.7 | GAAACAGGAGGTACACACAATGAGT | 83 | 94 | 96 | λ Rz1 homolog | |
| 35000–36760 | | 19 | TCAAGTAAGGAGGCAACGTGTCT | 586 | 99 | 86 | Large terminase subunit | |
| 35646–35879 | | 19.2 | GGAACTTGAGGATAAACCGTGGGT | 77 | 84 | 51 | Unknown | Overlaps gene 19; also found in φYeO3-12 |
| 36183–36356 | | 19.3 | AGCTGGAGGTTTCCGTGATGGTT | 57 | 84 | 65 | Unknown | Overlaps gene 19; also found in φYeO3-12 |
| 36842–36864 | φOR | | | | | | Phage promoter | Change from consensus: −5G |
| 37019–37168 | | 19.5 | ATAAAAGGGAGGAGACTCATGTTC | 49 | 97 | 65 | Unknown | |
| 37408–37555 | TR | | | | | | Concatemer formation | Direct terminal repeat |

[a] Genes are defined by the first base of the initiation codon and the last base of the termination codon. Sites of *E. coli* promoters are defined by positions −35 and +1, and those of φA1122 promoters are defined by −17 and +6. The consensus φA1122 (T7) promoter sequence is 5′-TAATACGACTCACTATAGGGAGA. RNase III recognition sites are defined by the first and last base of the stem-loop recognition structure; transcription terminators are defined by the first base of the stem-loop and the major site of termination.

[b] Initiation codons are underlined, and Shine-Dalgarno sequences are double underlined.

[c] aa, amino acids.

contain the heptamer sequence CCTAAAG (18), are also conserved in the φA1122 genome.

Except for gene *13*, all predicted φA1122 genes have easily recognizable Shine-Dalgarno sequences that at a minimum are GAGG or GGAG, or, for gene *1.8*, AGGA (Table 1). Twenty-nine of the 51 predicted genes can make six base pairs with the 3′ end of the 16S rRNA. However, only three consecutive base pairs can form between the gene *13* RNA and the 16S rRNA (Table 1; Fig. 2). An alignment of the φA1122 and T7 sequences 100 nucleotides upstream and downstream of the Shine-Dalgarno sequence revealed only four differences. Two are noncoding and do not affect any known regulatory element, and one is the change in R*13* described above; the fourth is in the Shine-Dalgarno sequence itself. In T7 this sequence is the normal GGAGG, but in φA1122 it has changed to GGATG. The presence of the T residue has been confirmed by sequencing directly from the phage genome. T7 gene *13* is essential, and it is expected that φA1122 gene *13* is also essential; presumably sufficient gp13 is made to support phage growth, despite the poor Shine-Dalgarno sequence.

**Early proteins.** The *0.3-0.7* fusion protein is active in preventing restriction by type I restriction enzymes, even though the annotations of the *Y. pestis* genome sequences (15, 51) do not refer to a type I restriction enzyme. However, the fusion does not have the *0.7* activity necessary to grow on the *rpoC* mutant strain BR3 (7, 45, 63). Host range mutants of φA1122 grow in *E. coli* containing *Eco*KI or *Eco*BI but not when RNAP contains the β′E2258K change present in BR3 (Table 2). These properties of the phage are not surprising; it is known that only the first 94 residues of T7 gp0.3 are necessary for antirestriction activity (5), and the *0.3-0.7* fusion protein contains the N-terminal 108 amino acids of gp0.3. However, there are only 35 residues that correspond to the C terminus of the 359-residue T7 gp0.7. T7 gene *0.7* is essential for growth on BR3 (63). Inspection of the nucleotide sequences on each side of the φA1122 fusion and T7 DNA shows that recombination between repeated GAGGAA sequences at T7 positions 1243 and 2996 would result in the synthesis of a gp0.3-0.7 fusion protein exactly equivalent to that encoded by φA1122.

Deletions inactivating *0.7* function are commonly found in T7 strains that have been passaged extensively under laboratory conditions (64). Interestingly, the apparent deletion in φA1122 DNA has a 5′ flanking sequence similar to those of the T7 mutants H1 and H1S (16, 63, 66), which are the most frequently isolated deletions affecting gene *0.7*. The T7 gene *0.7* mutants H1 and H1S were both isolated on the type I restriction-positive strain *E. coli* B and recognized by their failure to grow on strain BR3 (63). Host range mutants of φA1122 have the same properties (Table 2).

The φA1122 gene *1* product is 98% identical to T7 RNAP (Table 1). The φA1122 RNAP contains asparagine at position 748, which in T7 RNAP is responsible for differentiating between T7 and T3 promoters (55). As expected, therefore, all φA1122 phage promoters resemble their T7 counterparts. φA1122 gp1.2 is 88% identical to its T7 homolog and contains a similar anti-dGTPase activity, as judged by the plating efficiency of φA1122 (Table 2) on strains containing the *optA1* mutation (which increases host dGTPase levels), rendering *1.2* essential (58).

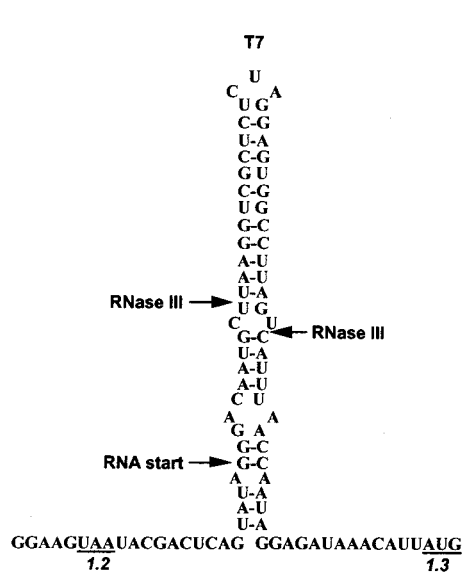T7 gene *1.2* is toxic in cells that express the F plasmid gene

*pifA*, and a *1.2* mutation is necessary for phage growth in the presence of F. The *1.2Y28N* change not only abolishes toxicity but is one of three mutations (the other two affect gene *10*) necessary for phage growth in the presence of F (40, 61). φA1122 gp1.2 contains glutamine at position 28. Consistent with the idea that tyrosine 28 in T7 gp1.2 is partially responsible for exclusion by F and that changes at that site help avoid exclusion, the relative plating efficiency of φA1122 host range mutants on isogenic *E. coli* strains with and without F is comparable to that of T7 *1.2Y28N* and significantly higher than that of wild-type T7 (Table 2).

**Nonessential late proteins.** φA1122 gp1.7 contains only 76 residues, about 100 fewer than its T7 or T3 counterparts. However, the predicted amino acid sequence of φA1122 gp1.7 is clearly related to the C-terminal residues of T7 and T3 gp1.7. The latter proteins also differ significantly in length, unlike most counterpart proteins from close T7 relatives. The function of gp1.7 is unknown, although a T7 *1.7* mutant has a reduced burst size (19).
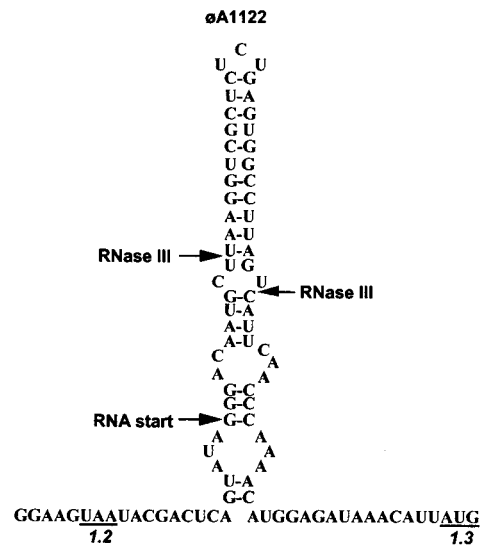
Positional counterparts to T7 genes *2.8* and *5.3* are missing in the φA1122 genome. Conversely, φA1122 contains two genes, *1.75* and *3.3*, that are not present in T7. Gene *1.75* overlaps the 3′ end of gene *1.7* by 20 bp and overlaps the 5′ end of the putative gene *1.8* by 50 bp. The putative gene *1.75* codes for an HNH endonuclease family protein commonly found in this group of phage genomes but displays no sequence similarity to any thus far described. The C-terminal two-thirds of gp1.75 is most closely related to gp179 of the mycobacteriophage Bzz1 and to an HNH endonuclease family protein in the *Streptococcus agalactiae* putative prophage λSa2. The predicted gene *3.3* product is yet another member of the HNH endonuclease family and, interestingly, shows sequence similarity to gp7.7 of both φA1122 and T7, to T7 gp2.8, and to gp5.3 of both T3 and φYeO3-12. Thus, the φA1122 genome contains two variants (gp3.3 and gp7.7) of the same group of HNH endonucleases. All of these putative endonucleases are likely expressed during phage infection, but it is not clear whether they play an important role in phage development. None of the counterpart T7 proteins is essential for growth, and no phenotype has yet been ascribed to their loss in mutant phages.

T7 gene *5.5* is normally nonessential, although it is required for growth on λ *rex*⁺ lysogens (36, 65). Although both the *Y. pestis* CO92 and KIM10+ genomes contain lambdoid and other prophage sequences (51, 15), annotations have not identified homologs to λ RexA or RexB. T7 gp5.5 also binds to the host DNA-binding protein H-NS, and an *E. coli hns* mutant that makes *5.5* essential has been isolated (37). *Y. pestis* contains *hns*, with the protein being about 85% identical to *E. coli* H-NS. φA1122 gp5.5 is 46 and 50% identical to, respectively, gp5.5 of T7 and φYeO3-12, but is 90% identical to T3 gp5.5. It is likely that the φA1122 and T3 genes *5.5* have the same activity, but whether T7 or φYeO3-12 gene *5.5* is directly equivalent is much less certain.
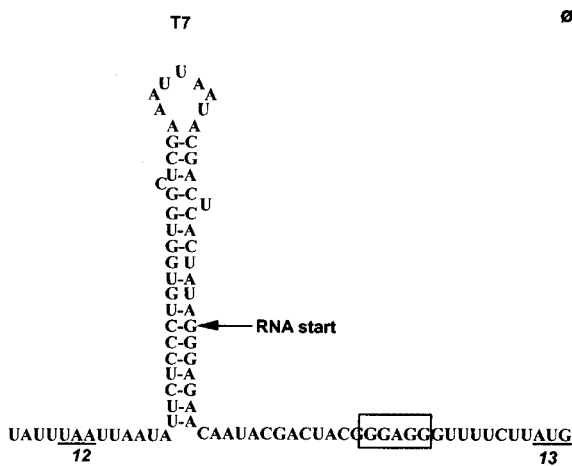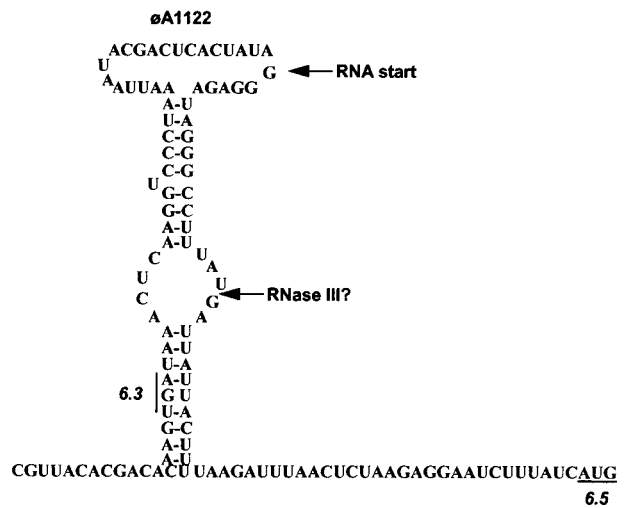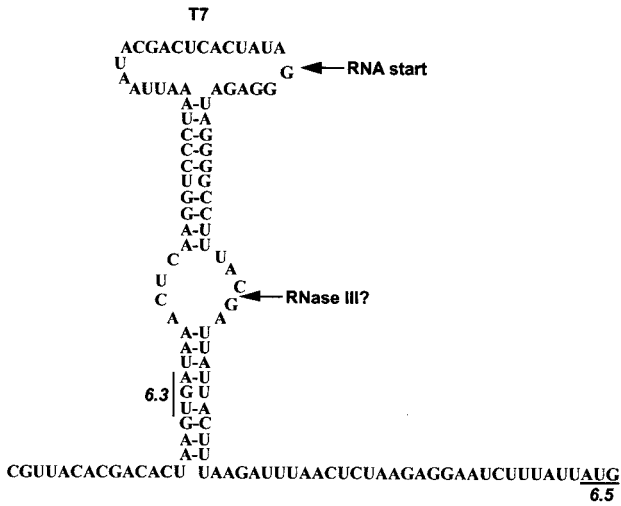
The φA1122 gene *3.5* product is 93% identical to its T7 counterpart, usually called T7 lysozyme. T7 lysozyme has two major functions: (i) inhibiting T7 RNAP activity from weaker promoters and increasing termination at the CJ terminator and (ii) facilitating lysis and release of progeny phage (38, 71, 73). The N-terminal residues of lysozyme that interact with RNAP

**ø1.3 R1.3**

T7

```
        U
     C     A
      U G
      C-G
      U-A
      C-G
      G U
      C-G
      U G
      G-C
      G-C
      A-U
      A-U
      U-A
RNase III → U G
      C   U ← RNase III
      G-C
      U-A
      A-U
      A-U
      C U
      A    A
       G A
      G-C
RNA start → G-C
      A   A
      U-A
      A-U
      U-A
```

GGAAG<u>UA</u>AUACGACUCAG GGAGAUAAACAUU<u>AUG</u>
1.2                              1.3

øA1122

```
        C
     U     U
      C-G
      U-A
      C-G
      G U
      C-G
      U G
      G-C
      G-C
      A-U
      A-U
RNase III → U-A
      U G
      C   U ← RNase III
      G-C
      U-A
      A-U
      A-U   C
      C      A
      A    A
      G-C
      G-C
RNA start → G-C
      A    A
    U      A
    A    A
      U-A
      G-C
```

GGAAG<u>UA</u>AUACGACUCA AUGGAGAUAAACAUU<u>AUG</u>
1.2                                1.3

**ø6.5 R6.5**

T7

```
ACGACUCACUAUA
U                G ← RNA start
 A AUUAA  AGAGG
      A-U
      U-A
      C-G
      C-G
      C-G
      U G
      G-C
      G-C
      A-U
      A-U
      C    U
      U     C
      C      C
      C     G ← RNase III?
      A    A
      A-U
      A-U
      U-A
  6.3 |A-U
      |G U
      |U-A
      G-C
      A-U
      A-U
```

CGUUACACGACACU UAAGAUUUAACUCUAAGAGGAAUCUUUAUU<u>AUG</u>
6.5

øA1122

```
ACGACUCACUAUA
U                G ← RNA start
 A AUUAA  AGAGG
      A-U
      U-A
      C-G
      C-G
      C-G
      U G-C
      G-C
      A-U
      A-U
      C    U
      U     U
      C      A
      C      G ← RNase III?
      A    A
      A-U
      A-U
      U-A
  6.3 |A-U
      |G U
      |U-A
      G-C
      A-U
      A-U
```

CGUUACACGACACU UAAGAUUUAACUCUAAGAGGAAUCUUUAUC<u>AUG</u>
6.5

**ø13 R13**

T7

```
     U  U
   A     A A
  A   U    A
  A A
  G-C
  C-G
  U-A
 C G-C
  G-C U
  G-C
  U-A
  G-C
  G U
  U-A
  G U
  U-A
  C-G ← RNA start
  C-G
  C-G
  U-A
  C-G
  U-A
```

UAUU<u>UA</u>AUUAAUA CAAUACGACUAC C[GGAGG]GUUUUCUU<u>AUG</u>
12                                       13

øA1122

```
    A U U U A
   A        U
  A          A
  G          C
  A          G
    U      C A
     C G U
       G U
       G-C
       U-A
       G-C
       G U
       U-A
       G U
       U-A
       C-G ← RNA start
       C-G
       C-G
       U-A
       C-G
       U-A
       U-A
       G-C
       U-A
```

UAUU<u>UA</u>AUUGAA AUACGACUACG[GGAUGG]UUUUCUU<u>AUG</u>
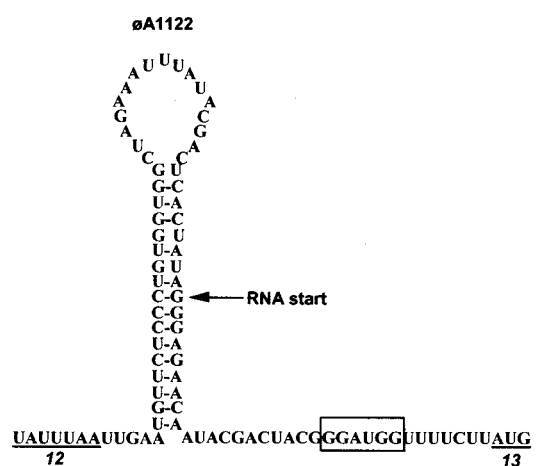12                                    13

FIG. 2. Predicted structures of three RNase III recognition sites in the φA1122 and T7 genomes. The termination codon of the upstream gene and the initiation codon of the downstream gene are indicated by underlines. The Shine-Dalgarno sequences for gene *13* are boxed. The T7 structures are adapted from reference 18, and the φA1122 RNase III was simply modeled to those structures for ease of comparison.

TABLE 2. Efficiencies of plating

| Bacterial strain | Relevant characteristics | Relative plating efficiency at 30°C[a] of phage strain[b]: | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | φA1122 | E6 | E11 | K18 | T3 | T7 | T7*1.2YN28* |
| A1122 | *Y. pestis* reference strain | 1 | 1.0 | 1.0 | 0.7 | $<10^{-8}$ | $7 \times 10^{-7}$ | ND[c] |
| IJ511 | *E. coli* K-12 reference strain | $5 \times 10^{-5}$ | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1 |
| IJ511/F*lac* | Pif$^+$ | ND | $5 \times 10^{-7}$ | $2 \times 10^{-7}$ | $4 \times 10^{-7}$ | 1.0 | $<10^{-10}$ | $2 \times 10^{-6}$ |
| IJ531 | IJ511 *optA1 tonA2* | ND | 1.0 | 1.1 | 1.2 | 1.0 | 1.1 | 1.2 |
| B | *E. coli* B prototroph | $4 \times 10^{-8}$ | 0.8 | 1.3 | 0.8 | 0.9 | 0.8 | 0.8 |
| BL2 | *E. coli* B *lig* | ND | 1.5 | 0.8 | 1.0 | 1.0 | 0.9 | 0.8 |
| BR3 | *E. coli* B *rpoC-E2258K* | ND | $6 \times 10^{-6}$ | $5 \times 10^{-6}$ | $4 \times 10^{-6}$ | 0.8 | 0.7 | 0.9 |
| BL21 | *E. coli* B Gal$^-$ *hsdS* | $10^{-6}$ | 1.3 | 1.3 | 1.0 | 1.0 | 1.1 | 1.1 |
| D$_2$371-48 | *S. sonnei* | ND | $<10^{-10}$ | $<10^{-10}$ | $<10^{-10}$ | 0.7 | $3 \times 10^{-7}$ | ND |

[a] Expressed relative to A1122 for φA1122 and relative to IJ511 for φA1122 host range mutants and both T3 and T7.
[b] The tail fiber protein of the host range mutant phage E6 contains residues 1 to 130 from T7, E11 contains T7 residues 118 to 125, and K18 is a spontaneous host range mutant containing the gp17-L523S change.
[c] ND, not determined.

(11) are fully conserved between φA1122 and T7, and as φA1122 RNAP is 98% identical to the T7 enzyme, it is probable that φA1122 gp3.5 has the same effects as T7 lysozyme on transcription during phage infection. Amino acid residues important for the amidase activity of T7 lysozyme are also conserved, and the φA1122 protein therefore should also contain this activity. Access of lysozyme to the cell wall requires the holin gp17.5; the φA1122 protein is 100 and 94% identical, respectively, to the T3 and T7 holins.

φA1122 contains homologs to the T7 overlapping genes *18.7, 19.2,* and *19.3*. Of these, gene *18.7* has an activity predicted to function during cell lysis when divalent cations are present (9). No function has been predicted for the other two genes, although T7 gene *19.3* is thought to be expressed (18).

**Essential late proteins.** The φA1122 essential protein that is least similar (76% identity) to its T7 counterpart is the tail protein gp7.3. With only 79 amino acids, the φA1122 protein is also considerably shorter than that in T7. T7 gene *7.3* is essential (Kemp and Molineux, unpublished data), although its function has not yet been clearly defined. The sequences of T3 and φYeO3-12 gp7.3 have also diverged from that of T7 gp7.3, sharing only 57 and 66% amino acid identity, respectively. However, the T7, T3, and φYeO3-12 proteins are comparable in length (106, 106, and 99 residues, respectively) The C-terminal 49 residues of T7 gp7.3 are identical to those of φA1122 gp7.3. Similarly, the C-terminal 33 residues of T3 and φYeO3-12 gp7.3 are 94% identical. Presumably it is the C termini of these proteins that contain the activity that renders the protein essential.

All other essential late proteins of T7 are more than 85% identical to their T7 counterparts, and it is thus reasonable to expect that all of the φA1122 homologs have equivalent properties (Table 1). For example, the active-form T7 DNA polymerase is a complex of gp5 and the host thioredoxin (39). φA1122 gp5 is 98% identical to T7 gp5; not surprisingly, therefore, φA1122 host range mutants fail to grow in *E. coli* lacking thioredoxin (data not shown). In addition, the φA1122 major capsid protein gp10 is 94% identical to T7 gp10, and in particular, residue 61 of the mature form of both proteins is glutamine. This residue can cause the abortive infection of T7 in *Shigella sonnei* D$_2$ 371-48 (29, 53). As expected, therefore, φA1122 host range mutants do not form plaques on *S. sonnei* D$_2$ 371-48 (Table 2), although they do adsorb (Kemp and

Molineux, unpublished observations). The efficiency of plating is, however, substantially lower than that of T7, suggesting that φA1122 may exhibit an additional growth defect in *S. sonnei* D$_2$ 371-48. However, we have no information on what that second defect may be. The abortive infection of T7 is due to endonucleolytic breakdown of the phage genome (27), a consequence of capsid protein synthesis in the infected cell (6). T7 DNA ligase is unable to prevent or repair the damage to the genome, although T3 DNA ligase possesses that capacity. One difference between T3 and T7 gp1.3 is a stretch of 19 amino acids that is missing in the T3 protein. However, φA1122 gp1.3 also contains a comparable "deletion"; thus, this region of nonhomology is unlikely to account for the difference in activity between the T3 and T7 enzymes in *S. sonnei* D$_2$ 371-48.

The primary determinant of phage host range is the tail fiber protein gp17. Both T3 and T7 tail fibers adsorb to the bacterial surface lipopolysaccharide, and the close similarity of φA1122 gp17 indicates that this phage also binds the surface lipopolysaccharide of *Y. pestis*. φA1122 infects only *Y. pestis*; the relative efficiency of plating on *E. coli* K-12 or B strains at temperatures of between 25 and 37°C is less than $10^{-4}$ (Table 2 and data not shown). An alignment of the gp17 proteins from φA1122 and T3 reveals that the φA1122 and T3 proteins are 98.9% identical; only 6 residues of 558 are different (Fig. 3). T3 does not plate on A1122 at detectable frequencies (Table 2); it does not adsorb efficiently. We have not yet isolated host range mutants of T3 that do grow. This is surprising, since one of the φA1122 mutants described below that grows almost equally well on *Y. pestis* and *E. coli* contains a tail fiber that has only five differences from T3 gp17. Perhaps two or more of these five residues must be changed before T3 can adsorb to *Y. pestis*. In contrast, T7 mutants that plate on *Y. pestis* arise at a frequency of about $10^{-6}$, even though the φA1122 and T7 proteins are only 86.4% identical, and include two short gaps necessary to maintain alignment.

**Host range mutants of φA1122.** φA1122 plates on the *E. coli* K-12 strain IJ511 at an efficiency of $5 \times 10^{-5}$ relative to the *Y. pestis* strain A1122 (Table 2). Plaques on IJ511 were picked and purified; none had switched host range specificity from *Y. pestis* to *E. coli*, and they had instead acquired an expanded host range that allowed growth on both bacterial species. All mutant phages plated on IJ511 at the same efficiency as on A1122. The gene 17 region of several independent mutants was

sequenced, and all but one (which contained two changes) were shown to contain a single mutation affecting the amino acid sequence of gp17 (E. Ramanculov, unpublished data). Although mutations were found throughout gene *17*, the mutant K18, containing the gp17L523S change, was chosen for further study described above. This change makes the amino acid at position 523 the same as T3 gp17 and is near the C terminus of the protein, which is the region of the tail fiber that interacts with the bacterial cell surface (62).

Host range mutants of φA1122 were also obtained after recombination with the T7 gene *17* plasmid pAR3685. The plasmid was electroporated into *Y. pestis* A1122 by using a selection for Amp[r], and the resulting strain was used as a host for φA1122. The titers of the resulting lysate on A1122 and IJ511 were then determined. The titers were approximately equal, indicating that at least half of the progeny phage had recombined with the multicopy T7 gene *17* plasmid. Such a high frequency of phage-plasmid recombination is not uncommon with T7 and T7 plasmids. Recombinants that grew on *E. coli* strain IJ511 were purified on the same strain, and all of those analyzed were found to have exchanged the 5′ portion of φA1122 gene *17* with T7 DNA from plasmid pAR3685. The overall nucleotide similarity between φA1122 and T7 gene *17* is 85.3%, but the 5′ 500 nucleotides are 93.6% identical, whereas the 3′ 500 nucleotides are only 73.4% identical. Thus, the pattern of recombinants was not unexpected from nucleotide similarities, but it was surprising that recombination over only the 5′ portion of gene *17* could result in a change in phage host range. The N-terminal portion of the T7 gp17 tail fiber is thought to interact with the phage tail (62) and is well conserved in other members of the T7 family with different adsorption specificities (e.g., *Y. enterocolitica* phage φYeO3-12 [49] and *E. coli* K1 capsule-specific phage K1F [52]).

Two recombinants from this experiment were selected; these contained the largest and the smallest replacements of φA1122 DNA with that of T7. E11 contains T7 codons 1 through 190 derived from T7; E6 contains only T7 codons 118 through 125. E6 gp17 has only a single amino acid change from the wild type, affecting codon 118; E11 gp17 contains the same change at codon 118 but also contains four other amino acid replacements (Fig. 3). Both E6 and E11 have an expanded host range and form plaques on *E. coli* at an efficiency comparable to that on *Y. pestis* A1122 (Table 2). Interestingly, changing part of the N-terminal region of φA1122 gp17 to the same sequence as T7 gp17 results in the same phenotype as the single amino acid change L523S near the C terminus of φA1122 gp17.

**Nucleotide sequence similarities between φA1122 and coliphages T7 and T3.** When the sequences of genes that have no positional counterpart (indels) in the other phage are removed, an alignment of the φA1122 and T7 genomes shows that they share about 89% identity; those of φA1122 and T3 share about 73% identity (Fig. 1). The latter value is essentially the same as that for the T3 and T7 genomes, which are about 74% identical (50). Thus, at the level of complete genomes, φA1122 is much more closely related to T7 than it is to T3.

Only two small regions of the T7 genome show less than 80% nucleotide identity with φA1122; one is centered on the normally nonessential gene *5.5* (37, 65), and the other is centered on genes *18* and *18.5/18.7* (Fig. 1). The longest stretch of identity between the φA1122 and T7 genomes is 482 nucleo-

tides (containing parts of genes *7.7* and *8*), and there are 5 stretches of >200 identical nucleotides and 30 stretches of >100 identical nucleotides.

A similar comparison between the φA1122 and T3 genomes is, however, considerably more revealing. Although there are only 13 stretches of sequence identity longer than 100 nucleotides across the entire genome, 5 are longer than 900 nucleotides (Table 3). Significantly, these five, plus several shorter stretches, lie in a single contiguous region whose ends lie in genes *15* and *19* (Fig. 1). Almost one-quarter of the φA1122 genome, a stretch of 9,188 nucleotides that codes for about half of the morphogenetic functions, is 99.8% identical between the φA1122 and T3 genomes (φA1122 positions 26066 to 35253; T3 positions 26657 to 35844). This degree of nucleotide conservation between the φA1122 and T3 genomes is quite remarkable. It is most unlikely that primary stocks of φA1122 have been exposed to *E. coli* since 1943 (33), and they perhaps have not been exposed since the first probable isolation of the phage (2). Wild-type φA1122 does not adsorb efficiently to the *E. coli* K-12 and B strains that we have used (although Hausmann [28] refers to a coliphage named A1122); conversely, T3, a phage first defined by Demerec and Fano (14), does not adsorb efficiently to *Y. pestis* A1122 (Ramanculov, unpublished observations). The two phages have thus been isolated in different laboratory collections for more than half a century, and they have been propagated on different hosts and most likely under different environmental conditions. Nevertheless, the two phages have retained an extremely high degree of sequence identity over a quarter of their genomes.

It seems obvious that φA1122 and T3 must be related through a recombination event that included the gene *15*-gene *19* region, and it is tempting to consider φA1122 a simple T7-T3 hybrid phage that acquired a host range mutation. However, this is a biased point of view, which arises because T3 has been more intensively studied genetically and biochemically than φA1122 and thus is more easily thought of as a parent phage. It is more logical to propose that φA1122 is a parent and T3 is the recombinant. Supporting this line of reasoning are the observations that the T3 φ*17* promoter has T7 promoter sequence and specificity (57), and T3 φ*17* is not used during infection. In contrast, the φA1122 φ*17* promoter (and all other φA1122 promoters) has the specificity of the φA1122 RNAP, which is presumed to be the same as that of T7. Thus, φA1122 φ*17* is more closely related to the remainder of the φA1122 genome than T3 φ*17* is to the remainder of the T3 genome. It follows, therefore, that φA1122 would be a parent phage and T3 would be a recombinant. The φ*17* promoter is contained within the gene *15*-gene *19* region of φA1122 that we suggest became part of the T3 genome.

Assuming that a single recombination event occurred, the 287-bp region of identity within gene *15* and the 1,351-bp region ending in gene *19* was donated from φA1122 to the genome that became defined as T3. The G+C content of this region is 48.8%, which is significantly less than that of the remainder of the T3 genome (50.3%) and much closer to that of the entire φA1122 genome (48.3%) and of its host *Y. pestis* (47.6%). Aside from the changes required for a different host range specificity, the majority of the mismatches between T3 and φA1122 in the gene *15*-gene *19* region presumably repre-

```
        ←────────────────────────────────────────────────────────────────────
T7      •••••••••• •••••••••• •••••••••• •••••••••• •••••••T•• •••••••••• ••••••••••
øA1122  MANVIKTVLT YQLDGSNRDF NIPFEYLARK FVVVTLIGVD RKVLTINADY RFATRTTISL TKAWGPADGY   70
T3      •••••••••• •••••••••• •••••••••• •••••••••• •••••••••• •••••••••• ••••••••••

                                                          ←──────→ (øA1122 E6)
T7      •••••••••• •••••••••• •••••••••• •••••••••• •••••••T•• •••••••••• ••••••••••
øA1122  TTIELRRVTS TTDRLVDFTD GSILRAYDLN VAQIQTMHVA EEARDLTADT IGVNNDGHLD ARGRRIVNLA  140
T3      •••••••••• •••••••••• •••••••••• •••••••••• •••••••••• •••••••••• ••••••••••

        ───────────────────────────────────────────────────→ (øA1122 E11)
T7      •••••••••• •••••••••• •••••••••• •••••••••• •A•G•••••S •••••••••• ••TK•F••••
øA1122  NAVDDRDAVP FGQLKTMNQN SWQARNEALQ FRNEAETFRN QTEVFKNESG TNATNTKQWR DEANGSRDEA  210
T3      •••••••••• •••••••••• •••••••••• •••••••••• •••••••••• •••••••••• ••••••••••

T7      KR•••••••• ••••••••S• •••••••••• ••••S••••H ••••Q••••• ••••••E•Y• ••••••••K••
øA1122  EQFKNTAGQY ATSAGNSATA AHQSEVNAEN SATAAANSAN LAEQHADRAE READKLGNFN GLAGAIDRVD  280
T3      •••••••••• •••••••••• •••••••••• •••••••••• •••••••••• •••••••••• ••••••••••

T7      ••••••••N• •••••••••T N••••••••• ••••VTN••• •••••••••• •••••••••• ••••G•••V
øA1122  GTNVYWKGGI HANGRLYLTS DGFDCGQYQQ FFGGSAGRYS VMEWGDENGW LMYVQRREWT TAIGDNIQLL  350
T3      •••••••••• •••••••••• •••••••••• •••••••••• •••••••••• ••H••••••• ••••••••V

T7      •••Q••T••• A•••Q••••• ••V•Q••••• •••H•••••• •••••••••• •••••••••• ••••H••T••
øA1122  VNGHIIAQGG DMTGPLKLQN GHALYLESAS DKAQYILSKD GNRNNWYIGR GSDNNNDCTF HSYVYGTNLT  420
T3      •••••••••• •••••••••• •••••••••• •••••••••• •••••••••• •••••••••• ••••••••••

T7      ••Q••••••• H••••••••• •••••••••• •••••••••R •SF•A•S••• •••••---S•• AG••V•V•--
øA1122  LKPDYAVVNK RFHVGQAVVA TDGNIQGTKW GGKWLDAYLN DTYVKKTKAW TQVWAAAGGS YMGGGSQTDT  490
T3      •••••••••• •••••••••• •••••••••• •••••••••• •••••••M•• •••••••S•• ••••••••••

T7      VS•••••••• •••CA••S•• •••••••••• •IASD••••R •••••••LG• •••A•SRSV• N••M••NE
øA1122  LPQDLRFRNI WIKTRNNYWN FFRTGPDGIY FLLAEGGWLK FQIHSNGRVF KNISDRDAPP TAIAVEDV    558
T3      •••••••••• •••••••••• •••••••••• ••S••••••• •••••••••• •••A•••••• ••••••••
                                      ↓
                                      S (øA1122 K18)
```

FIG. 3. Sequence alignment of the tail fiber protein gp17 of φA1122 with those of T7 and T3. Circles indicate identical residues; dashes in the T7 sequence were introduced to maintain alignment. Double-headed arrows represent the region of φA1122 gp17 that is replaced with T7 amino acids in the host range recombinants E6 and E11. The serine present in the spontaneous host range mutant K18 at residue 523 (gp17L523S) is indicated below the arrowhead.

sent adaptation of T3 to growth in *E. coli* (G+C content, 50.8%).

φA1122 may also have been the donor of a second region in the T3 genome. The products of φA1122 genes *5.5* through *5.9* are more similar to T3 than to T7 (Table 1), and a region encompassing the 3′ end of gene *5* (DNA polymerase) through gene *6* (exonuclease) has a higher level of sequence identity (Fig. 1). Inspection of the nucleotide sequences within this region reveals a 107-bp stretch of nucleotide identity in gene *5* and a 96-bp stretch within gene *6* (Table 2), which, ignoring the insertion of T3 gene *5.3*, provides a 3,156-bp segment that shows 95.0% nucleotide identity. This calculation assumes that T3 gene *5.3* entered the progenitor T3 genome after acquisition of φA1122 DNA. However, it is perhaps more likely that more than one recombination event occurred in this region. A genetic replacement of 1,772 bp may have occurred within DNA polymerase (gene *5*), using the 107-bp region of identity (Table 2) and a 77-bp stretch containing one mismatch (φA1122, 14277 to 14353; T3, 14915 to 14991) for recombination. This segment remains 97.1% identical in nucleotide se-

TABLE 3. Major regions of nucleotide conservation between φA1122 and T3

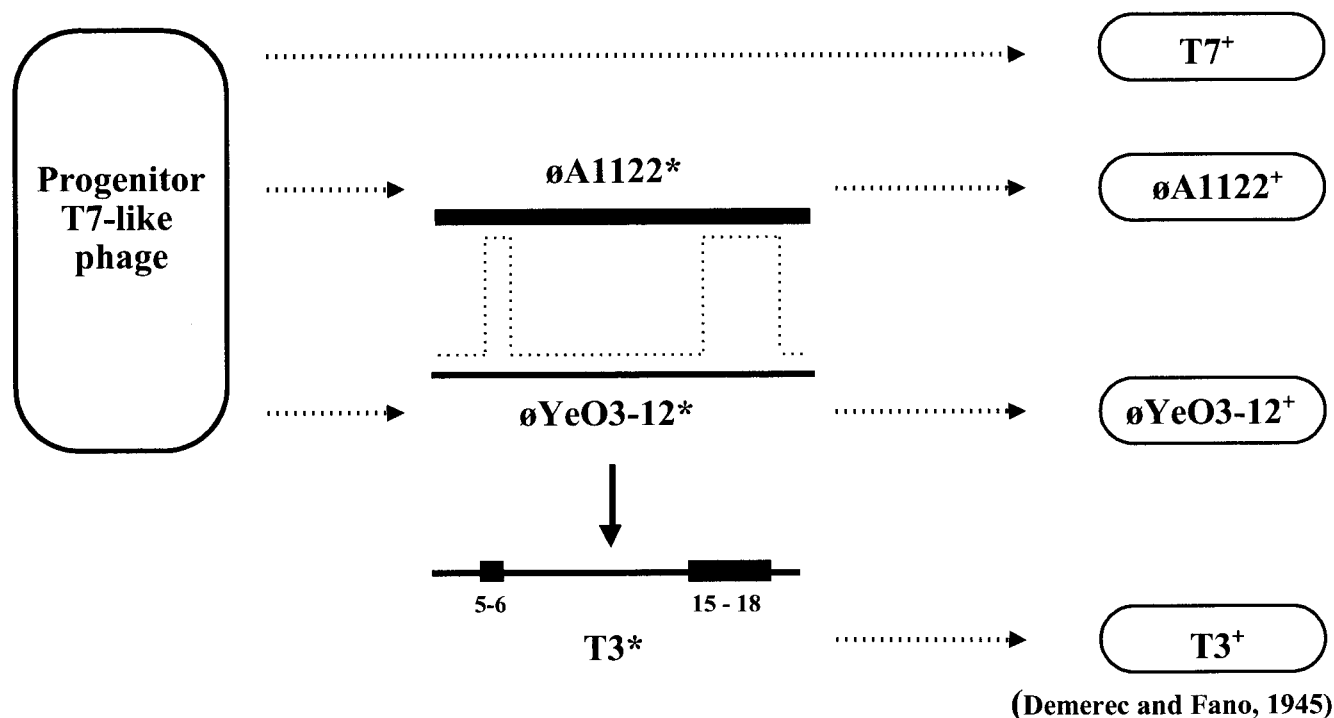| Nucleotide positions | | Length (nucleotides) | Gene |
|---|---|---|---|
| φA1122 | T3 | | |
| 28268–30200 | 28859–30791 | 1,933 | *16* |
| 30628–32318 | 31219–32909 | 1,691 | *16–17* |
| 26684–28266 | 27275–28857 | 1,583 | *15–16* |
| 33871–35221 | 34462–35812 | 1,351 | *18–19* |
| 32320–33236 | 32911–33827 | 917 | *17* |
| 33289–33642 | 33880–34233 | 354 | *17* |
| 26378–26682 | 26969–27273 | 305 | *15* |
| 30322–30626 | 30913–31217 | 305 | *16* |
| 26083–26369 | 26674–26690 | 287 | *15* |
| 13551–13693 | 14189–14331 | 143 | *5* |
| 33675–33807 | 34266–34398 | 133 | *17* |
| 30202–30320 | 30793–30911 | 119 | *16* |
| 12582–12688 | 13220–13326 | 107 | *5* |
| 13695–13791 | 14333–14429 | 97 | *5* |
| 15642–15737 | 16590–16685 | 96 | *6* |
| 14543–15636 | 15491–15584 | 94 | *5.5* |

FIG. 4. Possible ancestry of bacteriophage T3. The coliphage T7 and the yersiniophages φA1122 and φYeO3-12 are assumed to have evolved directly from a progenitor T7-like phage. Asterisks indicate the obvious uncertainty of the exact phages involved. φA1122∗ and φYeO3-12∗ are proposed to have recombined, probably in a *Yersinia* host. A recombinant phage subsequently underwent a switch in host range to yield what was defined by Demerec and Fano (14) as coliphage T3.

quence between φA1122 and T3. A second recombination event may have donated an 1,195-bp segment, now 95.1% identical, that spans the 5′ end of gene *5.5* through the 5′ half of gene *6*, using the 94- and 96-bp identical stretches (Table 3) for homology.

Ignoring regions of unequal genetic content, which correspond mainly to putative HNH or other endonucleases, the first 26 kb of the T3 genome, from the genetic left end through the beginning of gene *15*, is about 90% identical to that of the *Y. enterocolitica* phage φYeO3-12. Much of the remaining 12 kb of the T3 genome, from gene *15* through gene *19*, was found to be more similar to T7 (50). Consequently it was suggested that T3 may be the product of recombination between a T7-like coliphage and a YeO3-12-like yersiniophage. Indeed, recombination between T7 and the *Y. enterocolitica* phage φYeO3-12 in vivo was demonstrated experimentally; it allowed the isolation of a coliphage possessing many of the properties of T3 that distinguish it from T7. However, the level of sequence identity (91.0%) between the T3 and T7 genomes in the gene *15*-gene *19* region is significantly lower than that between the T3 and φA1122 genomes (98.4%) (also see Fig. 1). This suggests the alternative possibility that T3 is a host range variant of a recombinant between two yersiniophages. One phage is closely related to the *Y. enterocolitica* phage φYeO3-12, and the second is either the *Y. pestis* phage φA1122 itself or an extremely close relative (Fig. 4).

**A likely relationship between φA1122 and phage H.** φA1122 (Gunnison P phage) is used by the CDC as a diagnostic agent for *Y. pestis*. The U.S. Army still used phage H for similar

purposes. H was the name given by Molnar and Lawton (41) to the phage described by Cavanaugh and Quan (10), as there was no evidence that it and the Gunnison P phage were identical. Some properties of φA1122 and of phage H, as described by Cavanugh and Quan (10), are the same. Both are specific for *Y. pestis* at low temperatures but grow on *Y. pseudotuberculosis* at 37°C. However, at 37°C phage H attacks all strains of *Y. pestis*, about half of the *Y. pseudotuberculosis* strains, and all of the *E. coli* F⁻ strains tested (41). This host range is different from that of φA1122, which does not plate on *E. coli*, suggesting that phage H and φA1122 may be different. The host range of H, however, is comparable to those of the φA1122 host range mutants isolated as part of this study. Host range mutants of all members of the T7 phage family are readily isolated. For example, the original isolate of φII was described as specific for *E. coli* B (44, 69), but many widely distributed variants of φII plate normally on some K-12 strains (8, 42, 43, 64). Likewise it is possible that phage H is merely an expanded host range mutant of the original Gunnison P phage (which was renamed φA1122 for this study).

Heteroduplex mapping with phage H and φII DNAs (7) showed that the phages differ only in a single region, one that by its position likely carries gene *7.7*. The positions of nonhomologous regions in heteroduplexes of φII and T7 DNAs (30) are, with one exception, consistent with nonhomologous regions of the φA1122 and T7 genomes as shown here by nucleotide sequence comparison. The one exception is gene *7.7*, which is 99.7% identical between φA1122 and T7. Therefore, aside from gene *7.7*, the φA1122 and φII genomes are likely to be very similar. Since φII and phage H are so closely related

and since φA1122 was in use at the Hooper Foundation when phage H was said to be isolated, it seems likely that φA1122 and phage H were originally the same phage. Now that the genome sequence of φA1122 has been determined, and because this phage retains its absolute specificity for *Y. pestis* at temperatures below 28°C and for *Y. pestis* and *Y. pseudotuberculosis* at 37°C, we suggest that it should be used as the primary diagnostic phage in the identification of *Y. pestis* both in the United States and elsewhere.

## ACKNOWLEDGMENTS

## REFERENCES

1. **Adams, M. H.** 1959. The bacteriophages. Interscience Publishers, New York, N.Y.
2. **Advier, M.** 1933. Etude d'un bactériophage antipesteux. Bull. Soc. Pathol. Exotiques **26:**94–99.
3. **Altschul, S. F., W. Gish, W. Miller, E. W. Meyers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
4. **Arutiunov, I. I.** 1969. New modification of the increase of bacteriophage titer reaction (in plague). Lab Delo. **3:**179.
5. **Atanasiu, C., O. Byron, H. McMiken, S. S. Sturrock, and D. T. F. Dryden.** 2001. Characterisation of the structure of Ocr, the gene *0.3* protein of bacteriophage T7. Nucleic Acids Res. **29:**3059–3068.
6. **Beck, P. J., J. P. Condreay, and I. J. Molineux.** 1986. Expression of the unassembled capsid protein during infection of *Shigella sonnei* by bacteriophage T7 results in DNA damage that is repairable by bacteriophage T3, but not T7, DNA ligase. J. Bacteriol. **167:**251–256.
7. **Brunovskis, I., R. W. Hyman, and W. C. Summers.** 1973. *Pasteurella pestis* bacteriophage H and *Escherichia coli* bacteriophage φII are nearly identical. J. Virol. **11:**306–313.
8. **Buchstein, S. R., and D. C. Hinkle.** 1982. Genetic analysis of two bacterial RNA polymerase mutants that inhibit the growth of bacteriophage T7. Mol. Gen. Genet. **188:**211–218.
9. **Casjens, S., K. Eppler, R. Parr, and A. R. Poteete.** 1989. Nucleotide sequence of the bacteriophage P22 gene *19* to *3* region: identification of a new gene required for lysis. Virology **171:**588–598.
10. **Cavanaugh, D. C., and S. F. Quan.** 1953. Rapid identification of *Pasteurella pestis*: using specific bacteriophage lyophilized on strips of filter paper, a preliminary report. Am. J. Clin. Pathol. **23:**619–620.
11. **Cheng, X., X. Zhang, J. W. Pflugrath, and F. W. Studier.** 1994. The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. Proc. Natl. Acad. Sci. USA **91:**4034–4038.
12. **Chu, M. C.** 2000. Laboratory manual of plague diagnostic tests. CDC Press, Ft. Collins, Colo.
13. **Chung, Y.-B., C. Nardone, and D. C. Hinkle.** 1990. Bacteriophage T7 DNA packaging. III. A "hairpin" end formed on T7 DNA concatemers may be an intermediate in the processing reaction. J. Mol. Biol. **216:**939–948.
14. **Demerec, M., and U. Fano.** 1945. Bacteriophage-resistant mutants in *Escherichia coli*. Genetics **30:**119–136.
15. **Deng, W., et al.** 2002. Genome sequence of *Yersinia pestis* KIM. J. Bacteriol. **184:**4601–4611.
16. **Dunn, J. J., M. Elzinga, K. Mark, and F. W. Studier.** 1981. Amino acid sequence of the gene *0.3* protein of bacteriophage T7 and nucleotide sequence of its mRNA. J. Biol. Chem. **256:**2579–2585.
17. **Dunn, J. J., and F. W. Studier.** 1975. Effect of RNAse III cleavage on translation of bacteriophage T7 messenger RNAs. J. Mol. Biol. **99:**487–499.
18. **Dunn, J. J., and F. W. Studier.** 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of genetic elements. J. Mol. Biol. **166:**477–535.
19. **Endy, D., L. You, J. Yin, and I. J. Molineux.** 2000. Computation, prediction, and experimental tests of fitness for bacteriophage T7 mutants with permuted genomes. Proc. Natl. Acad. Sci. USA **97:**5375–5380.
20. **Ewing, B., and P. Green.** 1988. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. **8:**186–194.
21. **Ewing, B., L. Hillier, M. C. Wendl, and P. Green.** 1988. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8:**175–185.
22. **Flu, P. C.** 1927. Sur la nature du bactériophage. C. R. Soc. Biol. **96:**1148–1149.
23. **Girard, G.** 1943. Sensibilité des bacilles pesteux et pseudotuberculeux d'une part des germes du groupe coli-dysentérique d'autre part aux bactériophages homologues. Ann. Inst. Pasteur (Paris) **69:**52–54.
24. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
25. **Gunnison, J. B., A. Larson, and A. S. Lazarus.** 1951. Rapid differentiation between *Pasteurella pestis* and *Pasteurella pseudotuberculosis* by action of bacteriophage. J. Infect. Dis. **88:**254–255.
26. **Harrison, D. N., D. C. Cavanaugh, J. H. Rust, Jr., and J. D. Marshall, Jr.** 1971. Characteristics of a bacteriophage-infected strain of *Pasteurella pestis* isolated from a human case of plague. Infect. Immun. **4:**85–87.
27. **Hausmann, R.** 1968. Sedimentation analysis of phage T7-directed DNA synthesized in the presence of a dominant lethal gene. Biochem. Biophys. Res. Commun. **31:**609–615.
28. **Hausmann, R.** 1988. The T7 group, p. 259–289. *In* R. Calendar (ed.), The bacteriophages, vol. 1. Plenum Press, New York, N.Y.
29. **Hausmann, R., B. Gomez, and B. Moody.** 1968. Physiological and genetic aspects of abortive infection of a *Shigella sonnei* strain by coliphage T7. J. Virol. **2:**335–345.
30. **Hyman, R. W., I. Brunovskis, and W. C. Summers.** 1973. DNA base sequence homology between coliphages T7 and φII and between T3 and φII as determined by heteroduplex mapping in the electron microscope. J. Mol. Biol. **77:**189–196.
31. **Imburgio, D., M. Rong, K. Ma, and W. T. McAllister.** 2000. Studies of promoter recognition and start site selection by T7 RNA polymerase using a comprehensive collection of promoter variants. Biochemistry **39:**10419–10430.
32. **Inglesby, T. V., D. T. Dennis, D. A. Henderson, J. G. Bartlett, M. S. Ascher, E. Eitzen, A. D. Fine, A. M. Friedlander, J. Hauer, J. F. Koerner, M. Layton, J. McDade, M. T. Osterholm, T. O'Toole, G. Parker, T. M. Perl, P. K. Russell, M. Schoch-Spana, and K. Tonat.** 2000. Plague as a biological weapon: medical and public health management. JAMA **283:**2281–2290.
33. **Jawetz, E., and K. F. Meyer.** 1943. Avirulent strains of *Pasteurella pestis*. J. Infect. Dis. **73:**124–143.
34. **Larina, V. S., P. I. Anisimov, and A. K. Adamov.** 1970. A new plague phage strain for the diagnostics of the plague organism. Prob. Partic. Dangerous Infect. **11:**132–136.
35. **Lazarus, A. S., and J. B. Gunnison.** 1947. The action of *Pasteurella pestis* bacteriophage on strains of *Pasteurella*, *Salmonella*, and *Shigella*. J. Bacteriol. **53:**705–714.
36. **Lin, L.** 1992. Study of bacteriophage T7 gene *5.9* and gene *5.5*. Ph.D. dissertation. SUNY, Stonybrook, N.Y.
37. **Liu, Q., and C. C. Richardson.** 1993. Gene *5.5* protein of bacteriophage T7 inhibits the nucleoid protein H-NS of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **90:**1761–1765.
38. **Lyakhov, D. L., B. He, X. Zhang, F. W. Studier, J. J. Dunn, and W. T. McAllister.** 1998. Pausing and termination by bacteriophage T7 RNA polymerase. J. Mol. Biol. **280:**201–213.
39. **Mark, D. F., and C. C. Richardson.** 1976. *E. coli* thioredoxin: a subunit of bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **73:**780–784.
40. **Molineux, I. J., C. K. Schmitt, and J. P. Condreay.** 1989. Mutants of bacteriophage T7 that escape F restriction. J. Mol. Biol. **207:**563–574.
41. **Molnar, D. M., and W. D. Lawton.** 1969. *Pasteurella* bacteriophages sex-specific in *Escherichia coli*. J. Virol. **4:**896–900.
42. **Monner, D. A., and H. G. Boman.** 1970. Female strains of *Escherichia coli* K-12 as selective hosts for the isolation of female-specific mutants of phage φII. Biochem. Biophys. Res. Commun. **39:**1017–1020.
43. **Monner, D. A., S. Jonsson, and H. G. Boman.** 1971. Ampicillin-resistant mutants of *Escherichia coli* K-12 with lipopolysaccharide alterations affecting mating ability and susceptibility to sex-specific bacteriophages. J. Bacteriol. **107:**420–432.
44. **Monod, J., and E. Wollman.** 1947. L'inhibition de la croissance et de l'adaptation enzymatique chez les bactéries infectées par le bacteriophage. Ann. Inst. Pasteur (Paris) **73:**937–956.
45. **Nechaev, S., and K. Severinov.** 1999. Inhibition of *Escherichia coli* RNA polymerase by bacteriophage T7 gene *2* protein. J. Mol. Biol. **289:**815–826.
46. **Novoseltsev, N. N.** 1967. Studies on specificity and spectrum of action of plague moderate phage "H" and its virulent mutants. Genet. Biochem. Immunochem. Especially Dangerous Dis. **1:**88–98.
47. **Novosel'tsev, N. N., and V. I. Marchenkov.** 1990. *Yersinia pestis* phage of a new serovar. Z. Mikrobiol. Epidemiol. Immunobiol. **11:**9–12.
48. **Novosel'tsev, N. N., V. I. Marchenkov, and Y. Aurutiunov.** 1994. *Yersinia pestis* phages of serovar IV. Z. Mikrobiol. Epidemiol. Immunobiol. **6:**9–10.
49. **Pajunen, M. I., S. J. Kiljunen, M. E. Lotta Söderholm, and M. Skurnik.** 2001. Complete genomic sequence of the lytic bacteriophage φYe03–12 of *Yersinia enterocolitica* serotype O:3. J. Bacteriol. **183:**1928–1937.
50. **Pajunen, M. I., M. R. Elizondo, M. Skurnik, J. Kielecazwa, and I. J. Molineux.** 2002. Complete nucleotide sequence and likely recombinatorial origin of bacteriophage T3. J. Mol. Biol. **319:**1115–1132.

51. **Parkhill, J., et al.** 2001. Genome sequence of. Yersinia pestis, the causative agent of plague Nature **413:**523–527.
52. **Petter, J. G., and E. R. Vimr.** 1993. Complete nucleotide sequence of the bacteriophage K1F tail gene encoding endo-*N*-acylneuraminidase (Endo-*N*) and comparison to an Endo-*N* homolog in bacteriophage PK1E. J. Bacteriol. **175:**4354–4363.
53. **Pierce, J. C., and W. E. Masker.** 1988. A single base change in gene *10* of bacteriophage T7 permits growth on *Shigella sonnei*. J. Virol. **62:**4369–4371.
54. **Pokrovskaya, M. P.** 1940. Bacteriophage and its practical applications for treatment and prophylaxis of summer children's diarrhea, dysentery and surgical infections, p. 55–56. Kraipedizdat, Piatigorsk, Russia.
55. **Raskin, C. A., G. Diaz, K. Joho, and W. T. McAllister.** 1992. Substitution of a single bacteriophage T3 residue in bacteriophage T7 RNA polymerase at position 748 results in a switch in promoter specificity. J. Mol. Biol. **228:**506–515.
56. **Robertson, H. D., E. Dickson, and J. J. Dunn.** 1977. A nucleotide sequence from a ribonuclease III processing site in bacteriophage T7. Proc. Natl. Acad. Sci. USA **74:**822–826.
57. **Rosa, M. D., and N. C. Andrews.** 1981. Phage T3 contains an exact copy of the 23 base-pair phage T7 RNA polymerase promoter sequence. J. Mol. Biol. **147:**41–53.
58. **Saito, H., and C. C. Richardson.** 1981. Genetic analysis of gene *1.2* of bacteriophage T7: isolation of a mutant of *Escherichia coli* unable to support the growth of T7 gene *1.2* mutants. J. Virol. **37:**343–351.
59. **Saito, H., and C. C. Richardson.** 1981. Processing of mRNA by ribonuclease III regulates expression of gene 1.2 of bacteriophage T7. Cell **27:**533–542.
60. **Saito, H., S. Tabor, F. Tamanoi, and C. C. Richardson.** 1980. Nucleotide sequence of the primary origin of bacteriophage T7 DNA replication: relationship to adjacent genes and regulatory elements. Proc. Natl. Acad. Sci. USA **77:**3917–3921.
61. **Schmitt, C. K., and I. J. Molineux.** 1991. Expression of gene *1.2* and gene *10* of bacteriophage T7 is lethal to F plasmid-containing *Escherichia coli*. J. Bacteriol. **173:**1536–1543.
62. **Steven, A. C., B. L. Trus, J. V. Maizel, M. Unser, D. A. Parry, J. S. Wall, J. F. Hainfeld, and F. W. Studier.** 1988. Molecular substructure of a viral receptor-recognition protein. The gp17 tail-fiber of bacteriophage T7. J. Mol. Biol. **200:**351–365.
63. **Studier, F. W.** 1973. Genetic analysis of non-essential bacteriophage T7 genes. J. Mol. Biol. **79:**227–236.
64. **Studier, F. W.** 1979. Relationships among different strains of T7 and among T7-related bacteriophages. Virology **95:**70–84.
65. **Studier, F. W.** 1981. Identification and mapping of five new genes in bacteriophage T7. J. Mol. Biol. **153:**493–502.
66. **Studier, F. W., A. H. Rosenberg, M. N. Simon, and J. J. Dunn.** 1979. Genetic and physical mapping in the early region of bacteriophage T7 DNA. J. Mol. Biol. **135:**917–937.
67. **Sugino, T.** 1932. On the bacteriophage against the plague bacillus. Kitasato Arch. Exp. Med. **9:**72–81.
68. **Summers, W. C.** 2001. Bacteriophage therapy. Annu. Rev. Microbiol. **55:** 437–451.
69. **Wollman, E.** 1947. Relations entre le pouvoir de synthetiser la proline et la resistance au bacteriophage chez des mutants d'*Escherichia coli*. Ann. Inst. Pasteur (Paris) **73:**348–363.
70. **Zhang, N., and R. Young.** 1999. Complementation and characterization of the nested *Rz* and *Rz1* reading frames in the genome of bacteriophage. Mol. Gen. Genet. **262:**659–667.
71. **Zhang, X.** 1995. T7 RNA polymerase and T7 lysozyme: genetic, biochemical and structural analysis of their interaction and multiple roles in T7 infection. Ph.D. dissertation SUNY, Stonybrook, N.Y.
72. **Zhang, X., and F. W. Studier.** 1995. Isolation of transcriptionally active mutants of T7 RNA polymerase that do not support phage growth. J. Mol. Biol. **250:**156–168.
73. **Zhang, X., and F. W. Studier.** 1997. Mechanism of inhibition of T7 RNA polymerase by T7 lysozyme. J. Mol. Biol. **269:**964–981.
74. **Zhilenkov, Y. L., V. M. Fomchenkov, I. A. Novikov, V. E. Sadomov, M. V. Oborotov, and T. A. Gremyakova.** 1999. Examining the interaction of phages with microorganisms by fluorometry and electron-orientation spectroscopy. Vestn. Ross. Akad. Med. Nauk. **12:**24–29.