

A homing endonuclease and the 50-nt ribosomal bypass sequence of phage T4 constitute a mobile DNA cassette

Richard P. Bonocora¹, Qinglu Zeng^{1,2}, Ethan V. Abel³, and David A. Shub⁴

Department of Biological Sciences, University at Albany, State University of New York, Albany, NY 12222

Edited* by Marlene Belfort, Wadsworth Center, Albany, NY, and approved August 23, 2011 (received for review May 25, 2011)

Since its initial description more than two decades ago, the ribosome bypass (or “hop”) sequence of phage T4 stands out as a uniquely extreme example of programmed translational frame-shifting. The gene for a DNA topoisomerase subunit of T4 has been split by a 1-kb insertion into two genes that retain topoisomerase function. A second 50-nt insertion, beginning with an in-phase stop codon, is inserted near the start of the newly created downstream gene 60. Instead of terminating at this stop codon, approximately half of the ribosomes skip 50 nucleotides and continue translation in a new reading frame. However, no functions, regulatory or otherwise, have been imputed for the truncated peptide that results from termination at codon 46 or for the bypass sequence itself. Moreover, how this unusual mRNA organization arose and why it is maintained have never been explained. We show here that a homing endonuclease (*MobA*) is encoded in the insertion that created gene 60, and the *mobA* gene together with the bypass sequence constitute a mobile DNA cassette. The bypass sequence provides protection against self-cleavage by the nuclease, whereas the nuclease promotes horizontal spread of the entire cassette to related bacteriophages. Group I introns frequently provide protection against self-cleavage by associated homing endonucleases. We present a scenario by which the bypass sequence, which is otherwise a unique genetic element, might have been derived from a degenerate group I intron.

bacteriophage gene structure | group I intron | horizontal gene transfer | ribosomal frameshifting | mobile genetic element

In most of the phages of the T4 superfamily, the large subunit of DNA topoisomerase is a single polypeptide encoded by gene 39. However, in phage T4 this gene is disrupted by a 1-kb insertion, creating a truncated gene 39 plus a new gene (gene 60) that encodes the remainder of the topoisomerase subunit (1, 2) (Fig. 1). Furthermore, the C terminus of the truncated gene 39 and the N terminus of the newly created gene 60 each contain additional amino acid residues (43 and 30, respectively) that were not present in the original gene 39 homologs. These two independently translated proteins interact with the small subunit (encoded by gene 52) to create an enzymatically active topoisomerase (4). A second, 50-nt insertion beginning with an in-phase stop codon, is inserted into the newly created downstream gene 60 (2). Instead of terminating at this stop codon, approximately half of the ribosomes skip 50 nucleotides and continue translation in a new reading frame (5). Features involved in this remarkable process include: the stop codon, the codon at which translation resumes, a portion of the pre-hop nascent peptide, a stem-loop at the start of the bypassed sequence, a Shine/Dalgarno-like sequence within the bypassed RNA, and the structure of the bypassed mRNA (5–8).

The genome sequence of phage T4 (GenBank accession no. NC_000866) indicated that the insertion into gene 39 contains two ORFs: an apparent H-N-H homing endonuclease pseudogene (*mobA*), whose inferred translation product terminates after only 37 codons (Fig. 2), and gene 60.1, which overlaps the start of gene 60 and encodes a basic protein of 126 amino acids

with no relatives in the databases. Interestingly, a screen of independently isolated phages related to T4 revealed several (5/37) having the same architecture in their large subunit topoisomerase genes. Although DNA sequences of the ~1-kb insertions were not provided, each newly created gene 60 contained a ribosome bypass sequence identical to that of T4 (9, 10).

Results and Discussion

MobA Is a Site-Specific Endonuclease That Nicks T2 DNA Close to the Insertion Site of the T4 Bypass Sequence. Thinking that one of the insertions initially identified by Repoila et al. (9) might encode an intact *mobA* gene, we sequenced the PCR products of the gene 39/60 regions of two of them (phages Pol and SKX, provided by Henry Krisch, Laboratory of Microbiology and Molecular Genetics, CNRS, Toulouse, France) and resequenced this region from phage T4. Surprisingly, the insertions in all three phages are almost identical, encoding a single ORF of 271 amino acids (including both the *mobA* H-N-H domain and *g60.1*) that is highly similar to the other phage T4 H-N-H family homing endonucleases[†] (Fig. 2). The SKX and Pol *mobA* sequences are identical; the only differences from the T4 sequence are a synonymous G-to-A third position substitution in codon 156 and an A-to-G substitution at the second position of codon 169, changing threonine to alanine.

Free-standing homing endonucleases (i.e., those not encoded within a group I intron or intein) cleave the DNA of close relatives lacking the endonuclease gene, usually in a gene adjacent to the site of endonuclease gene insertion (12–14), and are copied into the recipient genome via recombination-mediated DNA repair, a process that has been called “homing” (for recent reviews, see refs. 14 and 15). Therefore, we tested the protein product of the *mobA* gene for endonuclease activity on DNA from phage T2, which has an uninterrupted gene 39.

The *mobA* gene of phage T4 was expressed in vitro using cell-free extracts and also in vivo by induction of plasmid expression vector pMobA. Incubation of either protein preparation with a

Author contributions: R.P.B., Q.Z., and D.A.S. designed research; R.P.B. and Q.Z. performed research; E.V.A. contributed new reagents/analytic tools; R.P.B., Q.Z., and D.A.S. analyzed data; and R.P.B., Q.Z., and D.A.S. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[†]R.P.B. and Q.Z. contributed equally to this work.

²Present address: Civil and Environmental Engineering, Massachusetts Institute of Technology, 15 Vassar Street, Cambridge, MA 02139.

³Present address: Thomas Jefferson University, 233 South 10th St., Philadelphia, PA 19107.

⁴To whom correspondence should be addressed. E-mail: shub@albany.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1107633108/-DCSupplemental.

[†]After this work was completed, the phage T4 genomic sequence was redetermined (accession no. HM_137666) with *mobA* sequence identical to that presented here (11). The inaccuracies in accession NC_000866 were undoubtedly due to the common practice at that time of cloning genes in unregulated plasmids before sequencing. Genes like *mobA*, whose products are highly toxic to *E. coli*, were likely to have suffered inactivating mutations at the cloning step.

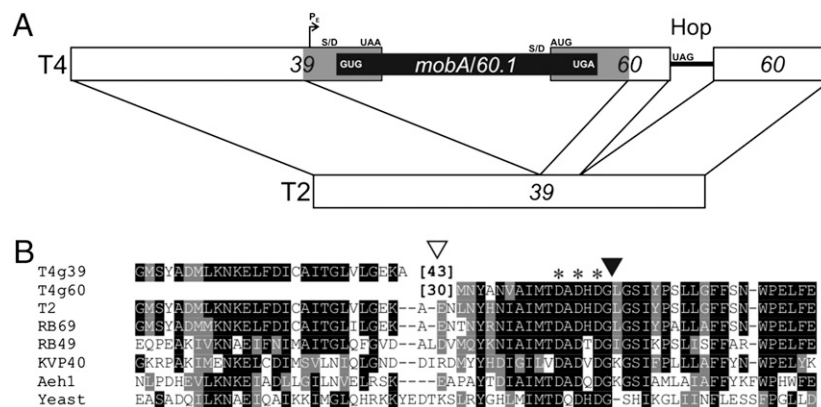


Fig. 1. Insertions in the phage T4 genes for the large subunit of topoisomerase II. (A) The parts of the T4 sequence that are homologous to phage T2 gene 39 are shown as open boxes. The carboxyl-terminal 43 amino acids of T4 gene 39 and the amino-terminal 30 amino acids of gene 60 are contributed by the insertion (gray bars), and the presumed *mobA* homing endonuclease pseudogene and gene 60.1 are depicted as a black bar. An early phage transcription start site (P_E), start and stop codons, and Shine/Dalgarno (S/D) sequences are indicated. The 50-nt untranslated bypass or Hop sequence (solid line) inserted after codon 46 of gene 60 begins with an in-frame stop codon. The schematic is not drawn to scale. (B) Amino acid sequences surrounding sites of insertion. The position of the 1-kb insertion containing *mobA* is indicated by an open triangle, with the number of additional amino acid residues added to the end of truncated gene 39 and the beginning of newly created gene 60 indicated in brackets. The insertion of the 50-bp bypass sequence is indicated by a solid triangle, with an asterisk showing positions of conserved active site residues. Homologous sequences from other members of the T4 superfamily and *Saccharomyces cerevisiae* are shown for comparison. Sequences were aligned with ClustalW1.8 (3). Identities and similarities of amino acid residues are outlined with Boxshade using default parameters (<http://sourceforge.net/projects/boxshade/>).

PCR-amplified region of T2 gene 39 generated a specific nick on the template (bottom) strand (Figs. 3 and 4B), but the proteins were inactive on the homologous T4 sequence (Fig. 4B). Introduction of a UAG stop codon at Tyr195 (Fig. 2) abolished activity (Fig. 3). To confirm that this activity was indeed encoded in T4 (and not in some other phage contaminant), we used DNA from a mutant T4 phage in our collection that has a deletion (*sa*Δ9) (16) located distantly from the *g39/60* locus as template for in vitro protein synthesis. This protein product displayed

similar nuclease activity on T2 DNA as when our wild-type T4 DNA was used to program the reaction (Fig. 3). Nucleotide resolution mapping of the cleavage position showed that MobA cleaves T2 DNA 19 bp upstream of the insertion site of the bypass sequence in T4 gene 60 (Fig. 4A and C).

Function of the Bypass Sequence Is Prevention of Self-Cleavage by MobA. The close association of the cleavage site and the bypass insertion site suggested that the bypass sequence might play

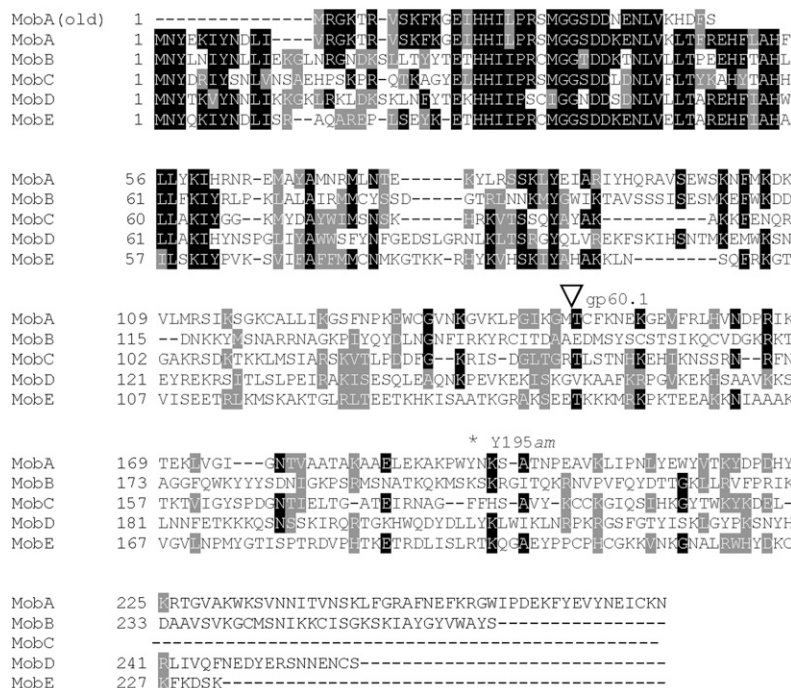


Fig. 2. Alignment of H-N-H family endonucleases of phage T4. Sequence of MobA (old) is as reported in accession no. NC_000866. All other sequences are as reported in T4 genomic sequence accession no. HM_137666. The beginning of putative gene product 60.1, within the MobA reading frame, is indicated by an open triangle. The location of the tyrosine to amber nonsense mutation at codon 195 is indicated by an asterisk. Alignment and shading were performed as in Fig. 1.

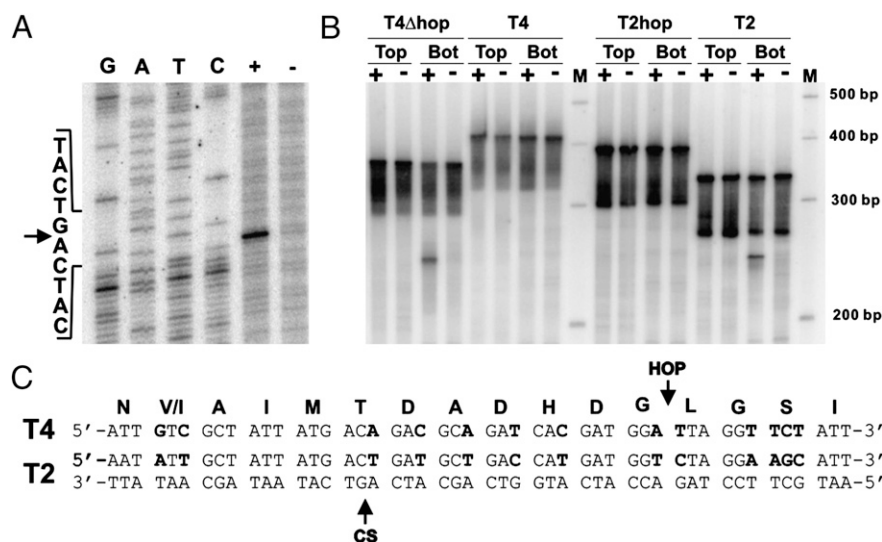


Fig. 4. MobA cleavage of phage DNA depends on the absence of the bypass sequence. (A) Determination of the cleavage site of MobA on T2 gene 39 DNA. Substrate DNA amplified with end-labeled bottom strand primer: (+) digested or (–) undigested with *in vitro*-synthesized MobA. The same end-labeled oligonucleotide was used to prime dideoxy sequencing reactions on cloned T2 DNA. All samples were subjected to electrophoresis in the same polyacrylamide gel. Sequence and position of cleavage (arrow) are shown on the left. (B) DNAs were amplified by PCR from plasmid templates, with either the top (Top) or bottom (Bot) strand primer end-labeled with 32 P. Amplified DNA was digested (+) or not digested (–) with MobA and subjected to denaturing gel electrophoresis next to molecular size markers (M). Target DNAs were from wild type T2 and T4, T4 with the bypass sequence deleted (T4Δhop), and T2 with the bypass sequence inserted (T2hop). (C) The DNA sequence of a portion of T2 gene 39 is shown, with the position of cleavage (CS) by MobA on the bottom strand indicated by an arrow. The coding strand of the homologous region of T4 gene 60 is shown above, with the position of insertion of the 50-nt bypass sequence (HOP) indicated by an arrow. Sequence differences are shown in boldface type. The translated amino acid sequence is shown above the nucleotide sequence.

implied by reduced translation efficiency, the bypass could have been maintained because of its essential relationship with *mobA*, with continued selection favoring increased efficiency of ribosome hopping. In this context, it is interesting to note that a portion of the nascent gp60 peptide chain upstream of the bypass sequence, which is required for efficient ribosome hopping (6, 7), is not homologous to a segment of the original gene 39, but is contributed as part of the 1-kb insertion that includes *mobA* (Fig. 14).

mobA is one example of a homing endonuclease-like gene that disrupts the coding sequence of an essential gene without destroying its function (1), and two other examples have recently been described in phages belonging to the T4 superfamily. In one case, a GIY-YIG family endonuclease gene disrupts DNA polymerase (28, 29); in the other case, a H-N-H family endonuclease gene disrupts the large subunit of ribonucleotide reductase (30). In all three cases, the products of the split genes reunite to form enzymatically active proteins. In the ribonucleotide reductase case, this requires peptide sequences that had been added to the ends of the split proteins at the site of the endonuclease gene

insertion (31). The C terminus of T4 gp39 and the N terminus of gp60 also have extensive amino acid sequences (43 and 30 residues, respectively) that are absent in the intact homologs (Fig. 1B), and it has been suggested (31) that these may also play a role in assembly of the split proteins into an active enzyme.

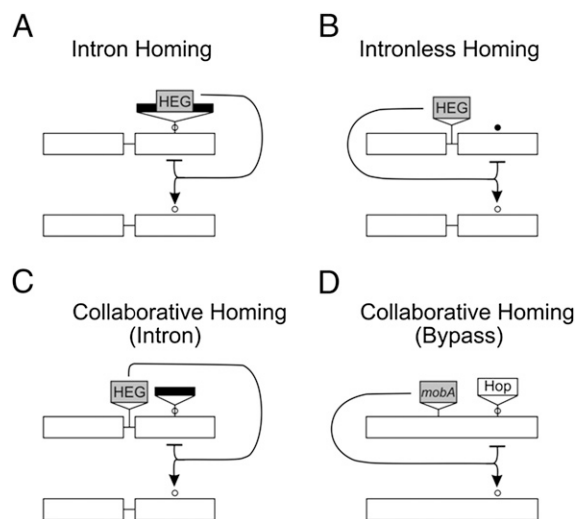


Fig. 5. Modes of protection against self-cleavage by homing endonucleases. (A) Homing endonuclease gene (HEG) encoded within a group I intron. The intron interrupts the sensitive binding/cleavage site (○). (B) HEG inserted intercistronically, resistance due to sequence differences at the binding/cleavage site (●). (C) HEG inserted intercistronically. Resistance is due to group I intron inserted into sensitive binding/cleavage site. (D) *mobA* inserted within an ancestral gene 39. Resistance is due to insertion of bypass sequence (Hop) within the sensitive binding/cleavage site. Group I introns, dark bars; homing endonuclease genes, shaded boxes; target genes without insertions, clear boxes; intergenic spaces, thin lines.

Table 1. Homing frequency

Expression plasmid*	Experiment	% L-arabinose			
		0	0.00002	0.0002	0.002
pMobA	1	97	99	94	56
	2	88	98	92	54
pF-CphI	1	1	0	2	0
	2	0	0	1	0
pBAD	1	2	1	2	0
	2	0	0	2	1

T2 progeny plaques (of 100 picked) that hybridize to the bypass sequence probe MobA25.

*Cells also contain the donor plasmid pT4 (39-60Δ*mobA*).

We have shown here that *mobA* can be spread by horizontal transmission, and it is likely that other similar constructs will also be shown to practice homing. Recent metagenomic analysis of environmental DNA samples has yielded additional examples of genes split by insertion of putative homing endonucleases (32), indicating that the infectious splitting of essential proteins may be a more common phenomenon than previously thought.

Materials and Methods

Oligonucleotides. Restriction sites are underlined in the sequences listed below.

ym001: AAATTAATACGACTCACTATAGGGATCCAAGCCATAGGGAGGACATATGAATTACGAAAAATCTATAATGA (noncoding sequence containing a T7 promoter and ribosome binding site is in italics);
ym003: GAGGATTTAATAGAACGACGAACATT;
M13Fwd: TGTAAACGACGGCCAGT;
M13Rev: CAGGAAACAGCTATGACC;
EVA7: GGGTCATGAATTACGAAAAATCTATAATGACTTAA;
EVA8: CCCGAATTCCTATATTTTTCATATCAACGCTAGAAGAA;
g39.5: CCCGAAACCAACGAGGATAAGTCATGA;
T2T4-3: AAAACTGCAGAGATTTTCCAAAGAGCCAAAGTCC;
T4dhop1: TCTGCGAGCTCATTATCTCTAGCGTGTGATATGAA;
T4dhop2: TCCGAGCAGAGAAGGATAAATAGAACCTAATCCATCGTGATCTGCGTCTGTC;
T4dhop4: GCTATTATGACAGACGAGATCACGATGGATTAGGTTCTATTTATCCTCTCTG;
T2hop1: TCTGCGAGCTCGATATTTGCGCAATCACTGGTCTA;
T2hop2: TCCAATAATCTCTTAATTAATGAGGTATTTCTATAGATAGCCGAAGGTAACCATCATGGTCAGCATCAGTCA (50-nt bypass sequence is in italics);
T2hop4: TAGCCTTCGGGCTATCTATAGAAATACCTCATAATTAAGAGATTATGGACTAGGAAGCATTTATCCTCTCTG (50-nt bypass sequence is in italics);
MobA6: GCTCTAGAAAGGTGACGATTATCTCCATGAGCAT;
MobA7: TATATAGGATCCCGCACTATTAAGTCATTATAGAT;
MobA8: TATATAGGATCCCAAGGAATGGTGCGGTGTTAATAAAG;
MobA10: TGAGCGGGATCCACAGGAGTTTGGACAAAGCGAATT;
MobA11: TTCTTTGGATCCCTGAGGGTGATTTCGGCTATCG;
MobA24: GTGAATTACGAAAAATCTATAATGAC; and
MobA25: TTCGGGCTATCTATAGAAATACCTCA.

Plasmid Construction. For expression of MobA, T4 *mobA* was amplified by PCR with KOD HiFi DNA polymerase (Novagen) using primers EVA7 and EVA8, digested with *PagI* and *EcoRI* and ligated into similarly digested pBAD/Myc-HisA vector (Invitrogen) to create pMobA. The 50-bp ribosomal bypass from phage T4 was inserted into a fragment of T2 gene 39 by the SOEing PCR method (33) with KOD HiFi DNA polymerase using primers T2T4-3, T2hop4, T2hop2, and T2hop1. The product was digested with *PstI* and *SacI* and ligated into pBSM13+ (Stratagene) to generate pT2hop. pT4dhop, which has the ribosomal bypass sequence removed from a fragment of the T4 gene 60, was created in a similar fashion using primers T2T4-3, T4dhop4, T4dhop2, and T4dhop1.

Plasmid pMobATS#4 was constructed by amplification of the T2 gene 39 with primers MobA10 and MobA11 and cloned in pBSM13+. pT4(39–60Δ*mobA*) was constructed as the donor plasmid for homing experiments. It contains all of gene 39 and all but the last seven codons of gene 60 of phage T4, to provide sequence homology for recombination with phage T2. To prevent the lethality that accompanies its uncontrolled expression, the plasmid also contains an in-frame deletion of *mobA* lacking codons 15–132: T4 gene 39 to codon 14 of *mobA* was amplified with primers g39.5 and

MobA7, and amplification from codon 133 of *mobA* to codon 153 of gene 60 was with primers MobA8 and MobA6. These products were digested with *BamHI* and ligated, and the purified product was digested with *PagI* and *XbaI* and ligated into similarly digested pACYC184.

In Vitro Expression of MobA Protein. Templates for in vitro protein synthesis were obtained by PCR amplification of T4 *mobA* with Taq DNA polymerase (Fermentas), according to the manufacturer's specifications, using primers ym001 and ym003. The purified product was used to direct MobA expression in vitro with TNT Quick Coupled Transcription/Translation Systems (Promega) as specified by the manufacturer.

In Vivo Expression of MobA Under Induction by L-Arabinose. A culture of *E. coli* LMG194 (Invitrogen) containing pMobA was grown to OD₆₀₀ = 0.5 in RM medium (1× M9 Salts, 2% Casamino Acids (Difco), 0.2% glucose, 1 mM MgCl₂) containing 100 μg/mL ampicillin, induced by adding L-arabinose to a final concentration of 0.02% at 37 °C for 5 h. Cells were harvested by centrifugation at 6,000 × g for 20 min, resuspended in equal volume ice-cold lysis mixture (50 mM Tris-HCl, pH 7.2, 1 mM EDTA, 1 mM PMSF, 2 μg/mL leupeptin, and 200 mM KCl), sonicated to complete lysis and used directly in endonuclease assays.

Generation of DNA Substrates. Individually 5' end-labeled targets were generated by 5' end labeling (with γ³²P]ATP and T4 polynucleotide kinase) one of the oligonucleotides before use in PCR together with an unlabeled primer partner (34). PCR was performed with Taq DNA polymerase as described above.

Endonuclease Assays. Endonuclease assay reactions (10 μL) containing 2 μL crude cell extract and 4 μL DNA substrate were incubated at 30 °C for 30 min in 0.05 M NaCl, 0.05 M Tris (pH 7.5), and 0.5 μg poly(dI-dC). Reaction products were extracted with an equal volume of phenol before separation on a 4% denaturing polyacrylamide gel.

Cleavage Site Mapping. The location of the MobA cleavage site was mapped, as previously described (34), on the bottom strand of the T2 gene 39 using substrate amplified from pMobATS#4 with labeled M13Fwd primer and unlabeled M13Rev primer. PCR amplification, 5' end labeling, and endonuclease assay conditions were the same as described above. Cleavage products were separated by electrophoresis on 4% denaturing polyacrylamide gel next to sequencing ladders generated using the same template DNA and labeled primers.

Plasmid to Phage Homing. Plasmid-to-phage homing assays were performed as described by Quirk et al. (18). Briefly, *E. coli* LMG194 cells harboring plasmids pT4(39–60Δ*mobA*) (donor) and either pMobA, pF-CphI, (35) or pBAD/Myc-HisA (the empty expression plasmid) were induced with increasing concentrations of L-arabinose on H agar plates at 30 °C for 2 h. Serial dilutions of T2 phage were spotted onto the plates and incubated overnight at 30 °C. The dilution that had almost cleared was scraped off, mixed with 1 mL buffered saline, shaken at 37 °C for 1 h, and diluted for single plaques on H plates. Plaques were transferred to a gridded lawn, and homing products were detected by plaque hybridization (17) using oligos MobA25 for the bypass sequence and MobA24 for the truncated *mobA* gene.

ACKNOWLEDGMENTS. This work was supported by the Northeast Biodefense Center through National Institutes of Health Grant U54-AI057158 and by a research development award from the College of Arts and Sciences, University at Albany.

- Huang WM, Wei LS, Casjens S (1985) Relationship between bacteriophage T4 and T6 DNA topoisomerases. T6 39-protein subunit is equivalent to the combined T4 39- and 60-protein subunits. *J Biol Chem* 260:8973–8977.
- Huang WM, et al. (1988) A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene 60. *Science* 239:1005–1012.
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.
- Seasholtz AF, Greenberg GR (1983) Identification of bacteriophage T4 gene 60 product and a role for this protein in DNA topoisomerase. *J Biol Chem* 258:1221–1226.
- Maldonado R, Herr AJ (1998) Efficiency of T4 gene 60 translational bypassing. *J Bacteriol* 180:1822–1830.
- Weiss RB, Huang WM, Dunn DM (1990) A nascent peptide is required for ribosomal bypass of the coding gap in bacteriophage T4 gene 60. *Cell* 62:117–126.
- Herr AJ, Nelson CC, Wills NM, Gesteland RF, Atkins JF (2001) Analysis of the roles of tRNA structure, ribosomal protein L9, and the bacteriophage T4 gene 60 bypassing signals during ribosome slippage on mRNA. *J Mol Biol* 309:1029–1048.
- Wills NM, et al. (2008) Translational bypassing without peptidyl-tRNA anticodon scanning of coding gap mRNA. *EMBO J* 27:2533–2544.
- Repoila F, Tétart F, Bouet JY, Krusch HM (1994) Genomic polymorphism in the T-even bacteriophages. *EMBO J* 13:4181–4192.
- F. Repoila (1995) Etude du Polymorphisme Genomique Chez les Bacteriophages T-pair. PhD thesis (Paul Sabatier University, Toulouse, France).
- Petrov VM, Ratnayaka S, Nolan JM, Miller ES, Karam JD (2010) Genomes of the T4-related bacteriophages as windows on microbial genome evolution. *Viral J* 7:292.
- Bonocora RP, Shub DA (2009) A likely pathway for formation of mobile group I introns. *Curr Biol* 19:223–228.
- Zeng Q, Bonocora RP, Shub DA (2009) A free-standing homing endonuclease targets an intron insertion site in the *psbA* gene of cyanophages. *Curr Biol* 19:218–222.

14. Edgell DR, Gibb EA, Belfort M (2010) Mobile DNA elements in T4 and related phages. *Virology* 7:290.
15. Raghavan R, Minnick MF (2009) Group I introns and inteins: Disparate origins but convergent parasitic strategies. *J Bacteriol* 191:6193–6202.
16. Depew RE, Snopek TJ, Cozzarelli NR (1975) Characterization of a new class of deletions of the D region of the bacteriophage T4 genome. *Virology* 64:144–145.
17. Belle A, Landthaler M, Shub DA (2002) Intronless homing: Site-specific endonuclease SegF of bacteriophage T4 mediates localized marker exclusion analogous to homing endonucleases of group I introns. *Genes Dev* 16:351–362.
18. Quirk SM, Bell-Pedersen D, Belfort M (1989) Intron mobility in the T-even phages: High frequency inheritance of group I introns promoted by intron open reading frames. *Cell* 56:455–465.
19. Landthaler M, Lau NC, Shub DA (2004) Group I intron homing in Bacillus phages SPO1 and SP82: A gene conversion event initiated by a nicking homing endonuclease. *J Bacteriol* 186:4307–4314.
20. Baranov PV, Fayet O, Hendrix RW, Atkins JF (2006) Recoding in bacteriophages and bacterial IS elements. *Trends Genet* 22:174–181.
21. Michel F, Westhof E (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol* 216:585–610.
22. Shub DA, et al. (1988) Structural conservation among three homologous introns of bacteriophage T4 and the group I introns of eukaryotes. *Proc Natl Acad Sci USA* 85:1151–1155.
23. Landthaler M, Begley U, Lau NC, Shub DA (2002) Two self-splicing group I introns in the ribonucleotide reductase large subunit gene of Staphylococcus aureus phage Twort. *Nucleic Acids Res* 30:1935–1943.
24. Lazarevic V (2001) Ribonucleotide reductase genes of Bacillus prophages: A refuge to introns and intein coding sequences. *Nucleic Acids Res* 29:3212–3218.
25. Swithers KS, Senejani AG, Fournier GP, Gogarten JP (2009) Conservation of intron and intein insertion sites: Implications for life histories of parasitic genetic elements. *BMC Evol Biol* 9:303.
26. Marchler-Bauer A, et al. (2009) CDD: Specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res* 37(Database issue):D205–D210.
27. Epstein RH, et al. (1963) Physiological studies of conditional lethal mutants of bacteriophage T4D. *Cold Spring Harb Symp Quant Biol* 28:375–392.
28. Petrov VM, et al. (2006) Plasticity of the gene functions for DNA replication in the T4-like phages. *J Mol Biol* 361:46–68.
29. Petrov VM, Ratnayaka S, Karam JD (2010) Genetic insertions and diversification of the PolB-type DNA polymerase (gp43) of T4-related phages. *J Mol Biol* 395:457–474.
30. Friedrich NC, et al. (2007) Insertion of a homing endonuclease creates a genes-in-pieces ribonucleotide reductase that retains function. *Proc Natl Acad Sci USA* 104:6176–6181.
31. Crona M, et al. (2011) Assembly of a fragmented ribonucleotide reductase by protein interaction domains derived from a mobile genetic element. *Nucleic Acids Res* 39:1381–1389.
32. Dassa B, London N, Stoddard BL, Schueler-Furman O, Pietrokovski S (2009) Fractured genes: A novel genomic arrangement involving new split inteins and a new homing endonuclease family. *Nucleic Acids Res* 37:2560–2573.
33. Horton RM, Cai ZL, Ho SN, Pease LR (1990) Gene splicing by overlap extension: Tailor-made genes using the polymerase chain reaction. *Biotechniques* 8:528–535.
34. Bonocora RP, Shub DA (2001) A novel group I intron-encoded endonuclease specific for the anticodon region of tRNA(fMet) genes. *Mol Microbiol* 39:1299–1306.
35. Zeng Q (2008) Collaborative homing directed by site-specific endonucleases in bacteriophages. PhD thesis (University at Albany, State University of New York, Albany, NY).

Supporting Information

Bonocora et al. 10.1073/pnas.1107633108

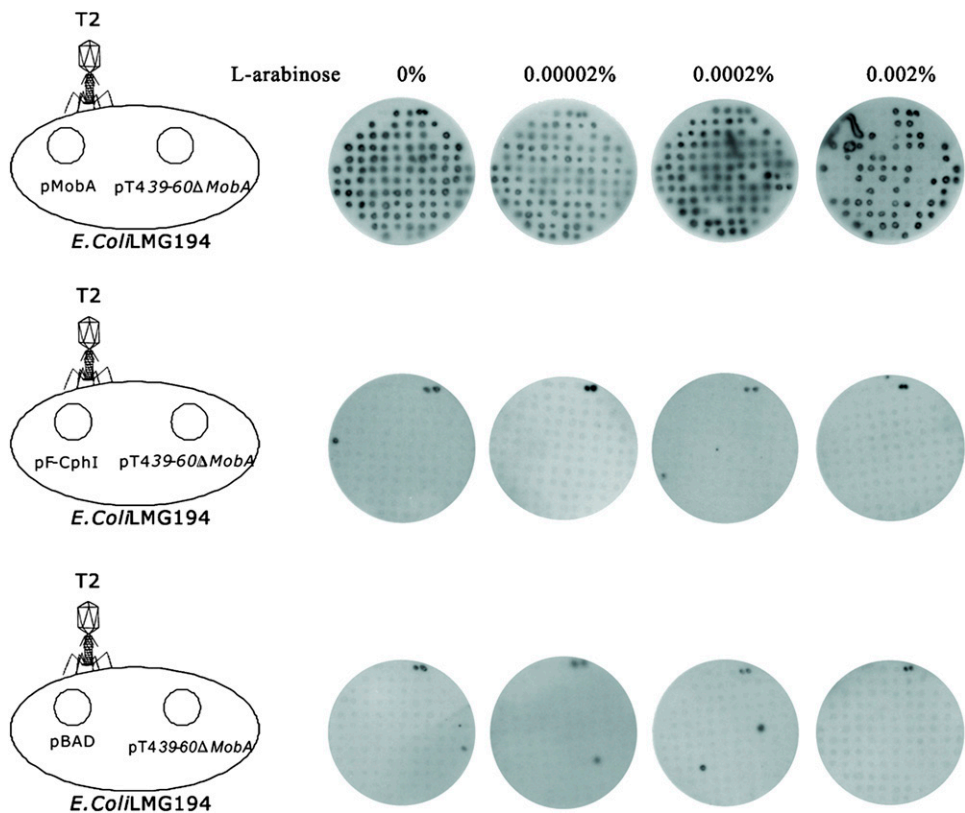


Fig. S1. MobA-dependent homing. Schematic representations of the two-plasmid homing experiment are shown on the left. *Escherichia coli* LMG194 cells harboring plasmid pT4(39–60ΔmobA) (donor) and different versions of expression vector pBAD were induced with various concentrations of L-arabinose and infected by phage T2. Plaques formed by progeny T2 were transferred to a fresh lawn with sterile toothpicks. Two wild-type T4 plaques were transferred at the upper right of each plate to act as positive hybridization controls. After development of clear zones, plaque lifts and DNA hybridization were carried out with oligonucleotide MobA25 to detect the presence of the bypass sequence. Quantitative data are presented in Table 1.

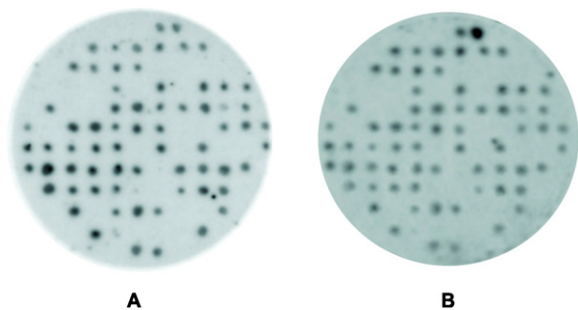


Fig. S2. Bypass sequence and truncated *mobA* gene are transferred together. *Escherichia coli* LMG194 cells harboring plasmids pT4(39–60ΔmobA) (donor) and pMobA were infected with T2 phage in the presence of 0.002% L-arabinose. Approximately half of the recovered T2 phages are expected to have acquired the bypass sequence (Table 1 and Fig. S1). One hundred progeny plaques were transferred to a fresh gridded lawn, with two wild-type T4 plaques transferred to the upper right to act as positive hybridization controls. After development of clear zones, plaque lifts were performed with two replica nylon membranes. Homing products were detected by hybridization using oligonucleotides MobA25 for the bypass sequence (A) and MobA24 for the 5' end of the truncated *mobA* gene (B).

Supporting Information

Bonocora et al. 10.1073/pnas.1107633108

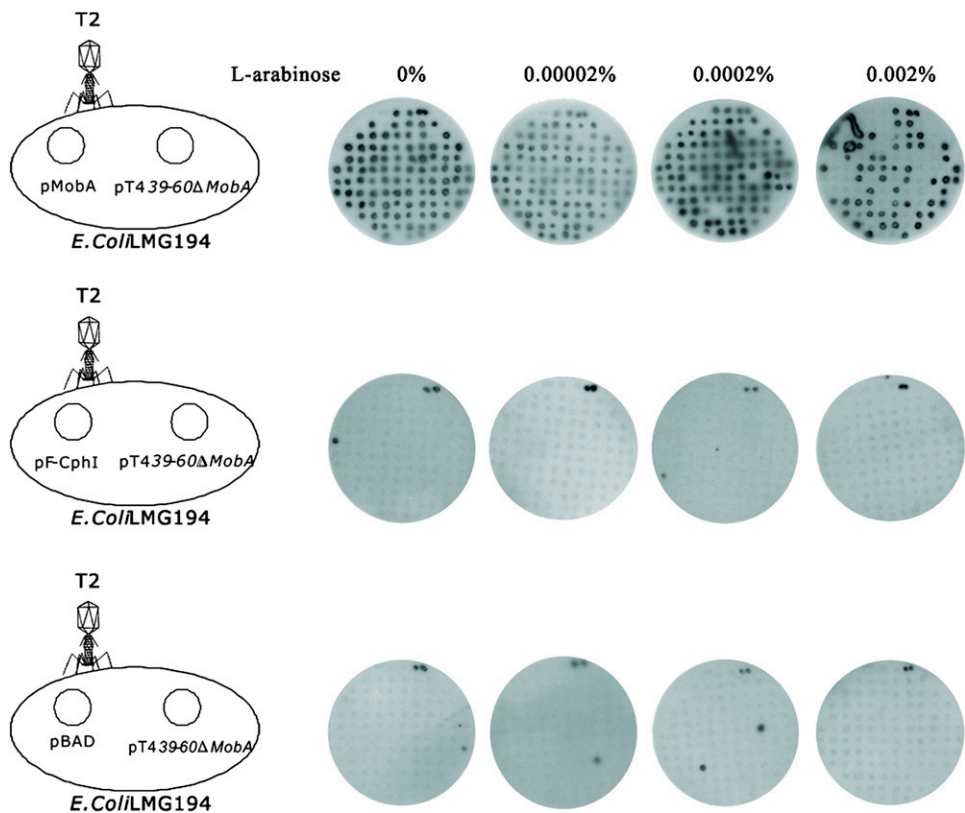


Fig. S1. MobA-dependent homing. Schematic representations of the two-plasmid homing experiment are shown on the left. *Escherichia coli* LMG194 cells harboring plasmid pT4(39–60ΔmobA) (donor) and different versions of expression vector pBAD were induced with various concentrations of L-arabinose and infected by phage T2. Plaques formed by progeny T2 were transferred to a fresh lawn with sterile toothpicks. Two wild-type T4 plaques were transferred at the upper right of each plate to act as positive hybridization controls. After development of clear zones, plaque lifts and DNA hybridization were carried out with oligonucleotide MobA25 to detect the presence of the bypass sequence. Quantitative data are presented in Table 1.

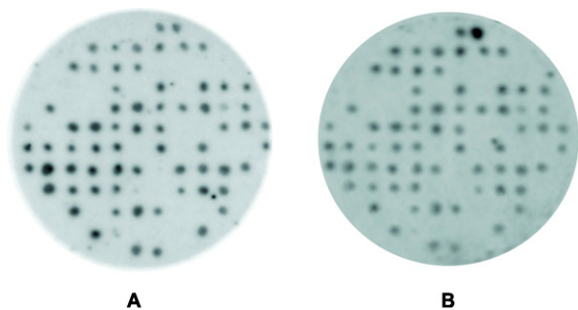


Fig. S2. Bypass sequence and truncated *mobA* gene are transferred together. *Escherichia coli* LMG194 cells harboring plasmids pT4(39–60ΔmobA) (donor) and pMobA were infected with T2 phage in the presence of 0.002% L-arabinose. Approximately half of the recovered T2 phages are expected to have acquired the bypass sequence (Table 1 and Fig. S1). One hundred progeny plaques were transferred to a fresh gridded lawn, with two wild-type T4 plaques transferred to the upper right to act as positive hybridization controls. After development of clear zones, plaque lifts were performed with two replica nylon membranes. Homing products were detected by hybridization using oligonucleotides MobA25 for the bypass sequence (A) and MobA24 for the 5' end of the truncated *mobA* gene (B).