



Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria

Michael Karberg, Huatao Guo, Jin Zhong, Robert Coon, Jiri Perutka, and Alan M. Lambowitz*

Mobile group II introns can be retargeted to insert into virtually any desired DNA target. Here we show that retargeted group II introns can be used for highly specific chromosomal gene disruption in *Escherichia coli* and other bacteria at frequencies of 0.1–22%. Furthermore, the introns can be used to introduce targeted chromosomal breaks, which can be repaired by transformation with a homologous DNA fragment, enabling the introduction of point mutations. Because of their wide host range, mobile group II introns should be useful for genetic engineering and functional genomics in a wide variety of bacteria.

Mobile group II introns—catalytic RNAs found in eubacteria and organelles—use a mobility mechanism termed retrohoming, which is mediated by a ribonucleoprotein (RNP) complex containing the intron-encoded reverse transcriptase (RT) and the excised intron lariat RNA (refs 1,2). The excised intron RNA inserts directly into one strand of a double-stranded DNA target site by a reverse splicing reaction, while the intron-encoded protein (IEP) site-specifically cleaves the opposite strand and uses the 3'-end of the cleaved strand for target DNA-primed reverse transcription (TPRT) of the inserted intron RNA. Several features of the retrohoming mechanism make it attractive for genetic manipulation. First, retrohoming is highly efficient and specific. Retrohoming frequencies can approach 100% and are six orders of magnitude higher than the frequency of ectopic integration^{1–6}. Second, the introns are minimally dependent on host factors. The TPRT mechanism requires only the IEP and the excised intron RNA (ref. 7), and the complementary DNA (cDNA) copy of the intron can be integrated into genomic DNA by recombination or repair mechanisms^{3–5}. The *Lactococcus lactis* Ll.LtrB intron, used as a model system, has a wide host range, being efficiently mobile both in its Gram-positive host and in Gram-negative *E. coli*⁴. Finally, unlike other integrating elements, group II introns can be retargeted to insert efficiently into virtually any desired DNA target⁸.

The key to retargeting group II introns lies in the mechanism by which they recognize their DNA target sites^{8–11}. The DNA target site for the Ll.LtrB intron extends from position –26 to +9 from the intron-insertion site, with a 15-nucleotide region from –12 to +3 recognized by base pairing of the intron RNA. The target-site sequences recognized by base pairing are denoted IBS1 and IBS2 (intron-binding sites 1 and 2) in the 5'-exon and δ' in the 3'-exon, and the complementary intron sequences are denoted EBS1, EBS2 (exon-binding sites 1 and 2) and δ (Fig. 1). The IEP recognizes only a few additional bases in the flanking regions of the DNA target site and facilitates local DNA unwinding, enabling the intron RNA to base pair to the IBS/ δ' sequences for reverse splicing^{9,11}.

Because most of the DNA target site is recognized by base pairing, the intron-insertion site can be controlled simply by modifying the intron RNA (refs 3,8–10). Previously, we developed an *E. coli* genetic system that enabled us to determine detailed target site-recognition rules for the *L. lactis* Ll.LtrB intron and to select introns that insert

into desired target sites from combinatorial intron libraries having randomized target site-recognition sequences⁸. We chose the HIV-1 provirus and the human gene encoding the HIV-1 co-receptor CCR5 as initial examples and showed that targeted group II introns could insert into these genes carried on plasmids in human cells⁸. Here we show that retargeted group II introns can be used for chromosomal gene disruption and modification in *E. coli* and two other bacteria.

Results

Selection of introns targeted to the *E. coli thyA* gene. We chose the *E. coli thyA* gene for initial experiments because chromosomal disruptants are scored readily by trimethoprim resistance (Tm^R)¹². Introns targeted to the *thyA* gene were selected from a combinatorial intron library using the previously developed, plasmid-based genetic system (Fig. 1A)⁸. In this system, a Δ -open reading frame (Δ ORF) derivative of the Ll.LtrB intron, carrying a phage T7 promoter near its 3'-end, inserts into a desired DNA target site cloned in a recipient plasmid upstream of a promoterless *tet^R* gene, thereby activating that gene. The intron-donor plasmid (pACD2) contains the modified intron and flanking exons cloned downstream of a T7lac promoter in a Cam^R vector, with the IEP (denoted LtrA protein) expressed from a position just downstream of the 3'-exon. Because the Δ ORF donor intron no longer encodes the LtrA protein, it cannot splice when it inserts at a new location.

To select group II introns targeted to *thyA*, we co-transformed the donor intron library (pDL2) into *E. coli* HMS174(DE3), with recipient plasmids pBRR3-*thyA*(s) and (a), which contain the *thyA* gene cloned in either the sense (s) or antisense (a) orientation upstream of the promoterless *tet^R* gene. Screening of Tet^R colonies obtained after induction of the host strain's isopropyl- β -D-thiogalactoside (IPTG)-inducible T7 RNA polymerase yielded 29 different introns that insert at 19 sites in the *thyA* gene. When recloned into the donor plasmid with modifications to correct mismatches in the EBS/IBS and δ/δ' pairings, the selected introns inserted into the cloned target sites at frequencies up to 2.6%, compared with 98% for the wild-type intron-target site combination (Fig. 1C and data not shown).

We tested six of the retargeted introns for their ability to insert into the chromosomal *thyA* gene (Fig. 1B,C). For chromosomal integration, donor plasmids containing the retargeted introns were

Institute for Cellular and Molecular Biology, Department of Chemistry and Biochemistry, and Section of Molecular Genetics and Microbiology, School of Biological Sciences, University of Texas at Austin, Austin, TX 78712. *Corresponding author (lambowitz@mail.utexas.edu).

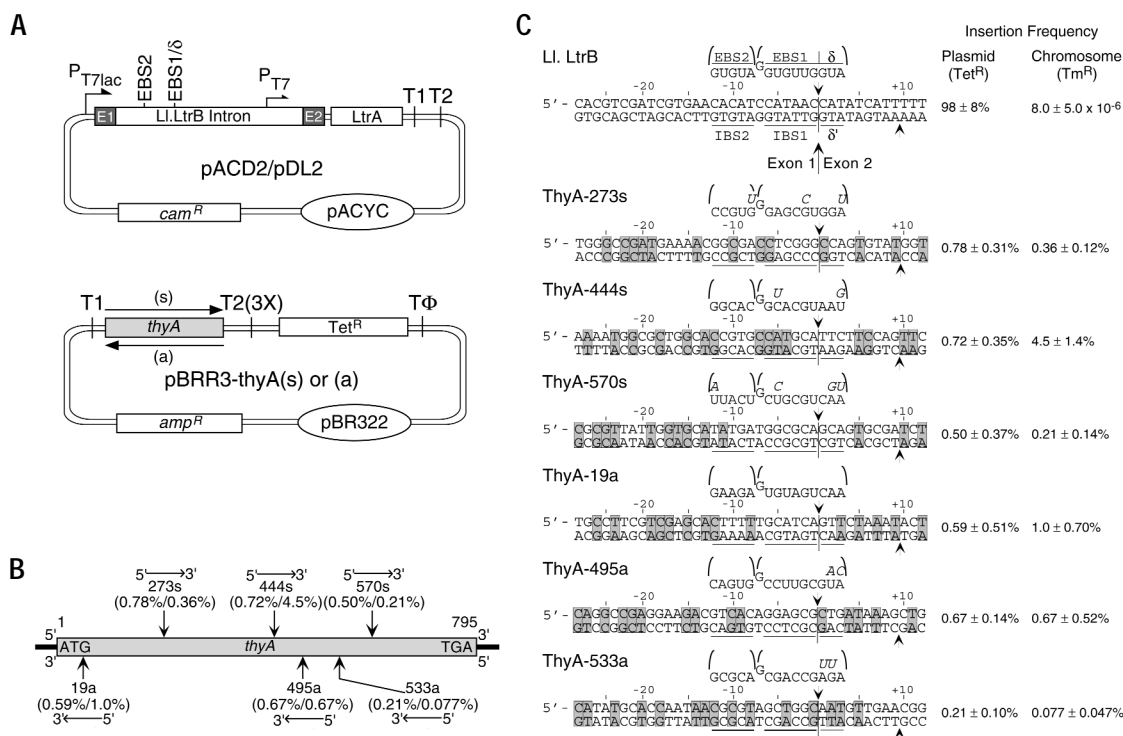


Figure 1. Group II introns targeted to the *E. coli thyA* gene. (A) Plasmid-based genetic system for the selection of targeted group II introns. The intron-donor plasmid pACD2 and donor intron library pDL2 (with randomized target site-recognition sequences (EBS2, EBS1, and δ)) contain a Δ ORF derivative of the LI.LtrB intron, with a phage T7 promoter (P_{T7}) inserted near its 3'-end. The intron and short stretches of flanking exons (E1 and E2) are cloned downstream of a T7lac promoter (P_{T7lac}) in a *Cam^R* pACYC184-based vector, with the LtrA ORF expressed from a position downstream of E2. The recipient plasmids pBRR3-*thyA*(s) and (a) contain the *E. coli thyA* gene cloned in the sense (s) or antisense (a) orientations upstream of a promoterless *tet^R* gene in an *Amp^R* pBR322-based vector. T1, T2, and T Φ are *E. coli mmbT1* and T2, and phage T7 T Φ transcription terminators. (B) Map showing insertion sites of anti-*thyA* introns in the *E. coli thyA* gene. Insertion sites in the top (sense) and bottom (antisense) strands are indicated by arrows above and below the gene, respectively. Introns are identified by the *thyA* position number of their insertion site (GenBank accession no. J01710), followed by "s" or "a" to denote orientation. The numbers in parentheses indicate plasmid/chromosome integration frequencies. (C) *thyA* target-site sequences and base-pairing interactions for selected anti-*thyA* introns. The wild-type LI.LtrB target site and base-pairing interactions are shown at the top. Nucleotide residues in the *thyA* target sites that match the wild-type target site are highlighted in gray. Modifications introduced into the EBS and δ sequences to optimize base pairing of the anti-*thyA* introns with their target sites are indicated above in italics. Intron-insertion frequencies are expressed as the percentage of Tet^R colonies in the plasmid assay and the percentage of Tm^R colonies for chromosomal integration. The insertion frequency of the wild-type LI.LtrB intron into a recipient plasmid containing the wild-type LI.LtrB target site (pBRR3-ltrB) was determined in parallel. Mobility frequencies are the average \pm s.d. for two to four independent experiments in each case.

transformed into *E. coli* HMS174(DE3) and induced with 100 μ M IPTG. Thy⁻ colonies were then selected by plating on minimal medium containing thymine and trimethoprim. The different anti-*thyA*

Table 1. Introduction of point mutations into the *E. coli thyA* gene by group II intron-stimulated homologous recombination^a

Transformants	Frequency
BL21(DE3) <i>recA</i> ⁺	3.9 × 10 ⁻⁶
+ Stop1	4.3 × 10 ⁻⁶
+ Stop2	5.7 × 10 ⁻⁶
+ ThyA-495a RT ⁻	1.1 × 10 ⁻⁵
+ Stop1 + ThyA-495a RT ⁻	5.3 × 10 ⁻⁵
+ Stop2 + ThyA-495a RT ⁻	5.6 × 10 ⁻⁵

^aCells were transformed with the indicated plasmids, induced with 100 μ M IPTG for 1 h at 37°C, grown in fresh LBT for an additional hour, and plated on minimal medium containing thymine (50 μ g/ml), with or without trimethoprim (20 μ g/ml). Plasmids pBR-*thyA*-Stop1 and 2 (Stop1 and 2) contain stop codons at different positions in the *thyA* gene (see Fig. 5). In the experiment shown, the frequencies of Tm^R colonies for cells transformed with both the group II intron and plasmid Stop1 or 2 were 5.3 and 5.6 × 10⁻⁵, ~15-fold higher than the spontaneous *thyA* mutation frequency of 3.9 × 10⁻⁶ in the absence of the intron or Stop1 or 2. In six repeats of the experiment, the homologous recombination frequencies in cells transformed with both the group II intron and Stop1 or 2 ranged from 1.1 to 7.2 × 10⁻⁵.

introns gave Tm^R colonies at frequencies of 0.077–4.5%, well above the spontaneous mutation frequency of 8.0 × 10⁻⁶ (Fig. 1C). Disruption of the chromosomal *thyA* gene was confirmed by PCR with flanking primers for at least 10 randomly selected colonies in each case, and by sequencing 10 independent events for the ThyA-444s and 495a introns (data not shown). In other experiments, with 500 μ M IPTG induction, ThyA-444s integrated at frequencies of 5–22%. By using appropriate mutants, we confirmed that chromosomal integration is dependent on both the RT and DNA endonuclease activities of the IEP, as expected for the TPRT mechanism (data not shown). Deletion of the internal T7 promoter had little if any effect on the chromosome integration frequency. Importantly, Southern hybridizations for the ThyA-444s and 495a disruptants showed that the retargeted introns inserted only at the desired chromosomal target site (Fig. 2A–C).

Disruption of other *E. coli* chromosomal genes. To test the generality of the approach, we designed group II introns targeted to different locations in four other *E. coli* genes: *lacZ*, *trpE*, *dadA*, and *proA*. The basic strategy was to identify the best matches to the "fixed" positions recognized by the IEP and then modify the intron to base pair to the adjacent 15-nucleotide region of the DNA target site for the EBS/BS and δ/δ' interactions^{8,10}. As shown in Figure 3, we readily obtained group II introns that insert into the desired chromoso-

mal target sites at frequencies of 0.1–10%, sufficient for detection by screening without selection. In each case, specific integration was confirmed by PCR and Southern hybridization (Fig. 2D and data not shown).

Generation of retargeted introns by PCR. For rapid chromosomal gene disruption, we developed a PCR-based method for generating the retargeted introns without cloning (Fig. 4). This method involves an initial two-step PCR with three unique primers and one “constant” primer to generate a 0.36-kilobase linear DNA containing the 5'-exon and 5'-end of the intron with modified IBS1/IBS2, EBS1/EBS2, and δ sequences. EBS1/EBS2 and δ are made complementary to the desired target site, while IBS1/IBS2 are made complementary to the modified EBS1/EBS2 for efficient RNA splicing (the δ/δ' interaction is not required for efficient splicing)⁸. The small linear DNA containing the modifications is then used as a megaprimer for a second PCR with a vector backbone containing the remainder of the donor intron in a thermosensitive plasmid with a *cam*^R marker. The resulting noncovalently closed, “circularized” PCR product is transformed directly into *E. coli*, where the nicks are sealed enabling replication as a plasmid. In a test with the ThyA-495a intron, the PCR product yielded chromosomal disruptants at a frequency of 0.20% compared to 0.45% for pACD-ThyA-495a. The slightly lower frequency for the PCR product likely reflects the lower copy number of the thermosensitive origin plasmid. By using this PCR strategy in conjunction with designed introns, it is possible to obtain essentially “overnight” chromosomal gene disruptions.

Use of retargeted group II introns in other bacteria. To test whether retargeted group II introns could be used in other bacteria, we co-transformed donor plasmids containing introns targeted to relatively conserved sequences in the *thyA* gene into *Shigella flexneri* and *Salmonella typhimurium* with plasmid pAR1219 (ref. 13), which encodes an IPTG-inducible T7 RNA polymerase. After induction with 100 μ M IPTG at 37°C for 1 h, Tm^R colonies were obtained at a frequency of 1% for pACD-ThyA-444s in *Shigella* and 0.3% for ThyA-19a in *Salmonella*, within the range of the integration frequencies in *E. coli*. Specific integration was again confirmed by PCR, sequencing, and Southern hybridization (data not shown).

Introduction of point mutations by group II intron-stimulated homologous recombination. Studies using rare-cutting endonucleases have shown that double-strand breaks in bacterial, plant, and animal cell chromosomal DNAs greatly stimulate homologous recombination with a co-transformed DNA fragment, enabling the introduction of desired mutations by gene replacement^{14–17}. In principle, group II introns can be used analogously with the advantage

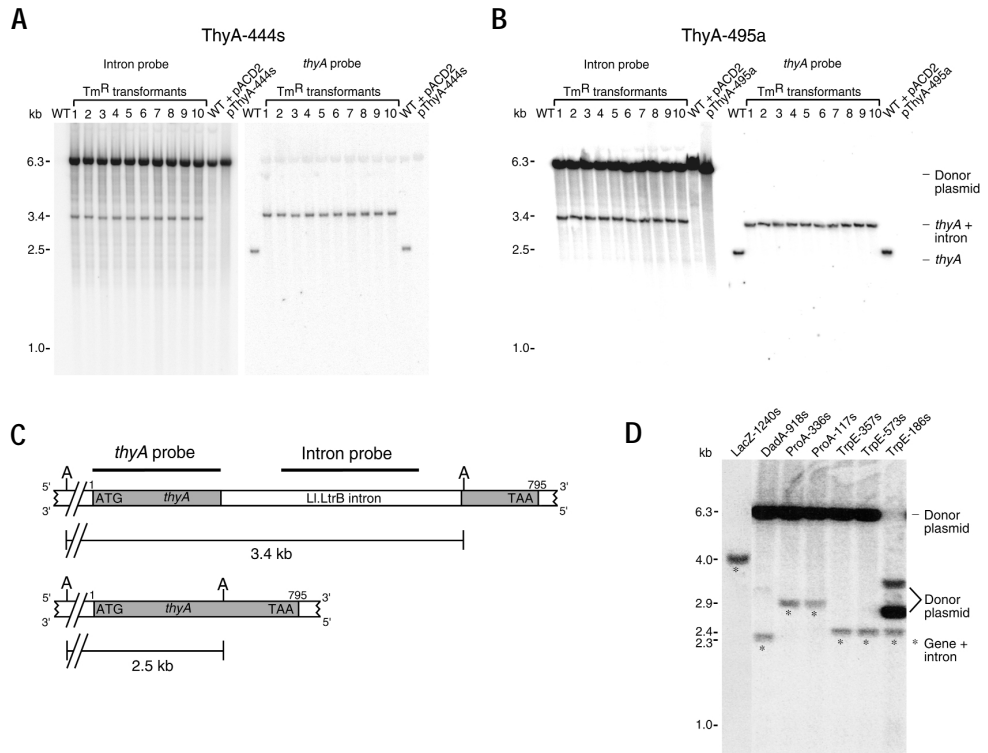


Figure 2. Southern hybridization showing disruption of *E. coli* chromosomal genes by targeted group II introns. (A, B) Southern blots for *thyA* disruptants. *Aa*I digests of DNAs from 10 independent Tm^R colonies (1–10) obtained with ThyA-444s-T7 or ThyA-495a-T7 (internal T7 promoter deleted by *Sa*I digestion followed by re-ligation) were hybridized with ³²P-labeled LI.LtrB intron (positions 294–935) or *thyA* (positions 1–519) probes. Control lanes show *Aa*I-digests of DNA from untransformed *E. coli* HMS174(DE3) (WT); *E. coli* HMS174(DE3) transformed with the untargeted wild-type LI.LtrB donor plasmid pACD2 (WT + pACD2); and the intron-donor plasmids pThyA-444s-T7 or pThyA-495a-T7. (C) Schematic showing restriction fragments detected by Southern hybridization for the *E. coli thyA* gene with and without the inserted intron. (D) Southern blots showing specificity of group II intron insertion into other *E. coli* genes. DNAs from randomly isolated disruptants obtained with the indicated retargeted group II introns in the experiment of Figure 3 were digested with *Fsp*I (*lacZ*) or *Fsp*I plus *Aa*I (*dadA*, *proA*, and *trpE*), which do not cut within the intron, and Southern blots were hybridized with the ³²P-labeled intron probe. The *lacZ* disruptant was plated on IPTG + X-Gal, which leads to loss of the donor plasmid. In the *TrpE*-186s disruptant, the modification of the IBS1 and IBS2 sequences in the donor plasmid introduces a *Fsp*I site, leading to two donor plasmid bands in the *Fsp*I/*Aa*I digest. These and all other disruptants analyzed (total >50) show a single intron insertion at the desired site. A, *Aa*I site.

that the double-strand break can be targeted to any desired region. Figure 5 outlines an experiment using this approach to introduce stop codons into the chromosomal *thyA* gene. In this experiment, we co-transformed a *recA*⁺ *E. coli* strain with an RT-deficient ThyA-495a intron and a cloned *thyA* template containing two TGA stop codons (asterisks) on either side of the ThyA-495a insertion site. The stop codons were either within (Stop1) or just outside (Stop2) the intron's recognition sequence. In the absence of cDNA synthesis by the RT-deficient IEP, the inserted intron RNA is apparently degraded, leaving a double-strand break at the target site. After induction with IPTG, cells transformed with both the targeted intron and the homologous fragment produced Thy⁻ colonies at frequencies of 5 to 6 × 10⁻⁵, ~15-fold higher than the spontaneous *thyA* mutation frequency (3.9 × 10⁻⁶) (Table 1). The introduction of the stop codons was confirmed by sequencing (see Fig. 5 legend for details). The stimulation of homologous recombination requires *recA* function and was relatively independent of the IPTG concentration used for induction (data not shown), suggesting that the introduction of the double-strand break was not limiting.

In another test, a chromosomal *thyA* gene disrupted by insertion of the wild-type LI.LtrB target site and a *cam*^R gene was repaired by co-transformation with the wild-type LI.LtrB intron and an overlapping wild-type *thyA* template. In this case, wild-type recombinants

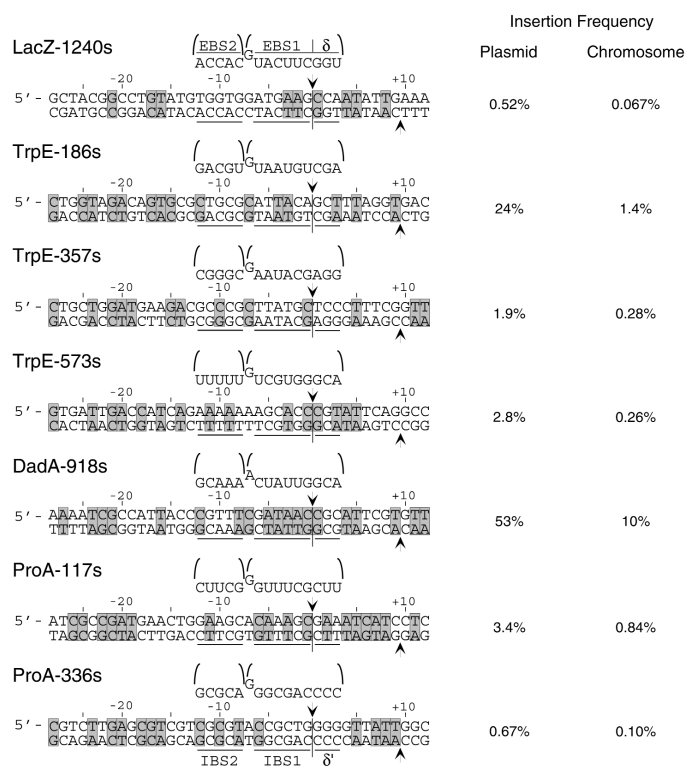


Figure 3. Group II introns targeted to *E. coli* chromosomal genes. Group II introns designed to insert into the *E. coli* genes *lacZ*, *trpE*, *dadA*, and *proA* were constructed by identifying the best matches to the “fixed” positions recognized by the IEP (refs 8,10) and then modifying the intron’s EBS2, EBS1, and δ sequences to base pair to target-site positions -12 to -8, -6 to -1, and +1 to +3, respectively, for the EBS/BS and δ/δ' interactions. In the *DadA*-918s intron, the RNA was modified to avoid a potentially deleterious GC pair at position -7 (ref. 8), and the +3 position was left unpaired. The designed introns were cloned into the pACD2 donor plasmid with complementary IBS1 and IBS2 sequences in the 5'-exon to insure efficient RNA splicing. The figure shows target sites and base-pairing interactions for the retargeted introns. Nucleotide residues that match the wild-type target site (Fig. 1C, top) are highlighted in gray. Insertion frequencies were determined for cloned target sites in the plasmid assay by Tet^R and for chromosomal gene disruption by blue/white selection on X-Gal (*lacZ*), auxotrophy (*trpE* and *proA*), or PCR (*dadA*). Similar frequencies were obtained in a repeat of the experiment for the LacZ-1240s, TrpE-186s, DadA-918s, and ProA-117s introns. The experiment was not repeated for the remaining three introns, which have relatively low mobility frequencies. Chromosomal insertion frequencies for ThyA-444s and 495a under the same induction conditions (500 μ M IPTG, 30°C) were 4.6 and 3.0%, respectively.

were obtained at a frequency of 10^{-7} , at least 100-fold higher than the recombination frequency for cells transformed with the wild-type *thyA* plasmid alone ($<10^{-9}$; not shown). The group II intron-based system should permit the introduction of any desired modification into any chromosomal gene, provided a suitable selection or screen is available.

Discussion

We show here that retargeted group II introns can be used for chromosomal gene disruption and modification in bacteria. The introns insert site-specifically into targeted *E. coli* genes at frequencies of 0.1–22% and, as demonstrated for *S. flexneri* and *S. typhimurium*, are readily adaptable to other bacteria simply by co-transforming the T7-driven intron-donor plasmid with a second plasmid encoding T7 RNA polymerase. The introns can also be expressed in other bacteria under the control of a native promoter, an approach taken to obtain chromosomal gene disruptions at frequencies $>10\%$ in the Gram-

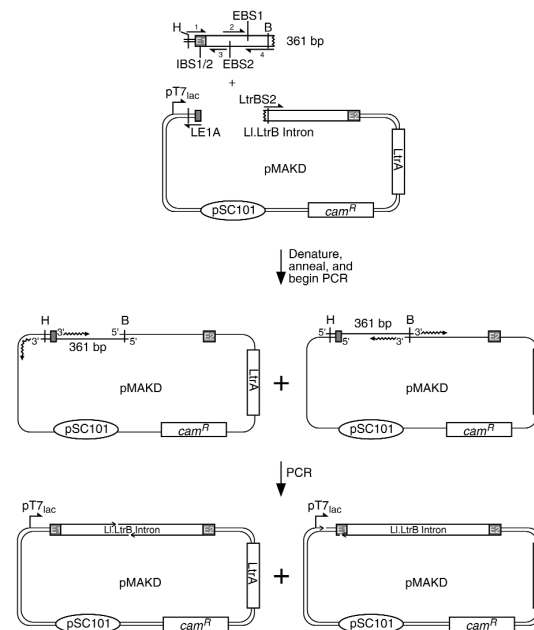


Figure 4. PCR strategy for generating retargeted group II introns without cloning. A 361-bp linear DNA (top) corresponding to the 5'-exon and 5'-end of the intron was generated in a two-step PCR by using two pairs of partially overlapping primers (primers 1 and 3, and primers 2 and 4) to introduce changes into IBS1/2 (primer 1), EBS2 (primer 2), and EBS1/ δ' (primer 4). Primer sequences were as follows: primer 1, 5'-AAAAAAGCTTCGTC GATCGTGAAGTCACAGGAGCGGTGC GCCCAGATAGGTTG; primer 2, 5'-CGCAAGTTTCTAATTCGTTGTGACTCGATAGAGAAAG TGCT; primer 3, 5'-AACCGAAATTAGAACTTGCGTTTCAG; primer 4, 5'-CAGATTGTACAAATGGTGTATAACAGATAAGTC GGAACGCTG AACTTACCTTTCTTTGT, where the underlined nucleotides are those changed for targeting to the *thyA* gene. The initial 361-bp linear DNA was then used as a megaprimer in a second PCR with the 8.2-kb vector backbone to produce a noncovalently closed, “circular” DNA, corresponding to a thermosensitive donor plasmid expressing the retargeted intron under the control of the T7lac promoter. In the experiment outlined, the vector backbone was generated by PCR of pMAKD, using primers LE1A (5'-TCACGATCGACAAGCTTTTTCTAGA GGGGAA) and LtrBS2 (5'-GTTATCACCACATTTGTACAATCTGTGG AGAA), but could also be generated by restriction digestion of a cloned plasmid. The final PCR product contains gaps (arrowheads) at different positions, depending on the strand of the 361-bp linear DNA from which priming occurred. After transformation into *E. coli*, the gaps are sealed by cellular enzymes to generate a replicating plasmid. The temperature-sensitive replication origin permits curing of the plasmid if desired. B, *BsrGI* site; H, *HindIII* site

positive bacterium *L. lactis* (D. Mills and A.M.L., unpublished data). For practical purposes, the donor introns can be designed by computation and generated by PCR without cloning, permitting essentially “overnight” disruptions. Because of the wide host range of group II introns, we anticipate that the procedures developed here will be directly applicable to a wide variety of bacteria.

Recently, alternative *E. coli* gene disruption systems have been developed that use phage λ Red or Rac RecE/RecT recombination functions to stimulate homologous recombination with linear PCR products or single-stranded oligonucleotides^{18–23}. The group II intron system is similarly facile in that the retargeted intron can be introduced as a PCR product. The efficiency of these systems differs for different genes. For group II introns, the frequencies range from 0.1–22% and were in almost all cases $>0.1\%$, at the higher end of the published values for the λ Red and ET systems (10^{-7} to 10^{-2}). With

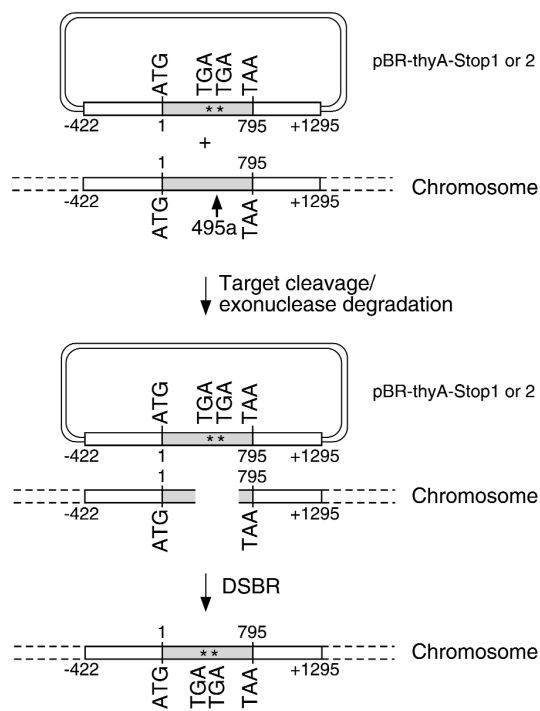


Figure 5. Introduction of point mutations into the wild-type *thyA* gene by group II intron-stimulated homologous recombination. *Escherichia coli* BL21(DE3) *recA*⁺ was transformed with a pACD2-based donor plasmid, which contains the ThyA-495a intron with an RT⁻ mutation, and pBR-thyA-Stop1 or 2, which contain the cloned *E. coli* HMS174(DE3) *thyA* gene, with stop codons (asterisks) at different positions in the *thyA* coding sequence. The *thyA* coding sequence is shaded, and the upstream and downstream sequences are unshaded and not drawn to scale. Frequencies of Tm^R colonies, indicative of loss of *thyA* function, are summarized in Table 1. The introduction of TGA stop codons was confirmed by DNA sequencing: For Stop1, 16/20 Tm^R colonies contained the TGA stop codons, with 13 containing both TGAs and 3 containing only the upstream TGA; for Stop 2, 20/22 Tm^R colonies contained the stop codons, with 13 containing both TGAs, 4 containing only the upstream TGA, and 3 containing only the downstream TGA. Three of the six remaining Tm^R colonies (2 for Stop1 and 1 for Stop2) had other mutations at or within 100 bp of the target site and may have been generated by incorrect double-strand break repair. Such mutations may also contribute to the background of Tm^R colonies for cells transformed with the ThyA-495a intron alone (see Table 1).

insertion frequencies >0.1%, it is possible to use group II introns in the absence of a selectable marker, making them more amenable to obtaining strains with multiple disruptions. Finally, with respect to use in other bacteria, group II introns have minimal dependence on host-specific factors, retain full activity in *recA*⁻ hosts (cf. ref. 22), and unlike λ Red, are not dependent on a highly efficient electroporation procedure for introduction of sufficiently high concentrations of DNA oligonucleotides.

In addition to the applications demonstrated here, group II introns with foreign genes inserted in domain IV can be used as vectors to integrate these genes at desired chromosomal locations (J.Z. and A.M.L., unpublished data)^{4,24}. In principle, group II introns can also be used to obtain conditional disruptions by linking splicing of an integrated Δ ORF intron to expression of the IEP from a separate, inducible promoter. In practice, however, we find that the splicing of the Ll.LtrB intron inserted into *E. coli* chromosomal genes is frequently inefficient, so that the possibility of obtaining conditional disruptions must be evaluated in each case.

Finally, in another application, we show that group II introns can

be used to introduce point mutations into chromosomal genes by targeting double-strand breaks that stimulate homologous recombination with a co-transformed DNA fragment containing the desired modification. Although the recombination frequencies obtained here using group II introns (10^{-4} to 10^{-5}) are lower than those reported for the λ Red system, they are still suitable for routine genetic manipulation. Furthermore, the group II intron system may be useful for stimulating homologous recombination in other bacteria that are less amenable to λ Red and for introducing targeted double-strand breaks in eukaryotes.

Experimental protocol

Bacterial strains. *E. coli* HMS174(DE3) *recA*⁻ (Novagen, Madison, WI) and BL21(DE3) *recA*⁺ (Stratagene, La Jolla, CA) were used for intron mobility experiments, and DH5 α and DH10B for cloning. Other bacteria were *S. flexneri* SA100 (ref. 25) and *S. typhimurium* JR501 (ref. 26). Antibiotics were ampicillin, 100 μ g/ml; chloramphenicol, 25 μ g/ml; tetracycline, 25 μ g/ml; and trimethoprim, 20 μ g/ml (50 μ g/ml for *S. typhimurium*).

Recombinant plasmids and libraries. The intron-donor plasmid pACD2 (Fig. 1A) was previously denoted pACD- Δ ORF+ORF2 (ref. 8). Derivatives of pACD2 include the RT-deficient mutant pACD2-RT⁻ (conserved YADD changed to YAAA)²⁴ and pMAKD in which a 4.6-kb *AseI/Bsu36I* fragment containing the temperature-sensitive replication origin of pMAK705 (ref. 27) replaces the 2.4-kilobase *AseI/Bsu36I* fragment containing the pACYC184 origin. The combinatorial intron library pDL2 was constructed from pACD2 by inserting random sequences at EBS2 positions -12 to -8, EBS1 -6 to -1, and δ +1 to +3, as well as the corresponding IBS2-IBS1 positions to provide complementary nucleotide combinations for RNA splicing⁸.

The recipient vector pBRR3 is a derivative of pBRR1 (previously denoted pBRR-Tet)⁸, which contains three (instead of one) *rrnB* T2 transcription terminators (116-bp *EcoRI/XbaI* fragment of pBRR1) inserted between the polylinker used for cloning the target DNA and the promoterless *tet^R* gene (Fig. 1A). The T2 sequences terminate *E. coli* but not T7 RNA polymerase and thus reduce background expression from cryptic *E. coli* promoters in cloned target DNAs. pBRR3-*thyA*(s) and (a) contain the *E. coli thyA* gene (801-bp *EcoRI/SphI* fragment of pSUThyA; ref. 10) cloned in the sense (s) or antisense (a) orientations upstream of the promoterless *tet^R* gene in the *XhoI* site of pBRR3.

pBR-thyA-Stop1 and 2 contain the cloned *thyA* gene of *E. coli* HMS174(DE3) (2.3 kb *HincII* fragment) beginning 422 bp upstream of the *thyA* coding sequence and ending 1,295 bp downstream, with stop codons introduced at different positions in the coding sequence (TAT₄₉₀₋₄₉₂TGA, TCC₄₉₉₋₅₀₁TGA in Stop1 and GCA₄₆₃₋₄₆₅TGA, TTC₅₂₆₋₅₂₈TGA in Stop 2). The stop codons are positioned on either side of the intron ThyA-495a insertion site, within (Stop1) or just outside (Stop2) of the intron's recognition sequence (positions 486-521).

Group II intron mobility assays and chromosomal gene disruption. Group II intron mobility and selection were carried out with intron-donor and recipient plasmids in *E. coli* HMS174(DE3)⁸. After 1 h induction with 100 μ M IPTG at 37°C (*thyA*) or 500 μ M IPTG at 30°C (*lacZ*, *trpE*, *proA*, and *dadA*), cells were plated on Luria-Bertani (LB) medium containing ampicillin, with or without tetracycline, and mobility events were selected as Tet^R colonies. For chromosomal gene disruptions, cells were transformed with a donor plasmid containing the targeted intron and induced with IPTG as above.

PCR method for generating retargeted group II introns. A 361-bp PCR product corresponding to the 5'-exon and 5'-end of the intron (E1-25 to I+326) was generated by an initial two-step PCR, using two pairs of partially overlapping primers, three of which introduce modifications into the IBS1/2, EBS2, and EBS1/ δ sequences (see Fig. 4). The gel-purified 361-bp PCR product was then used in a second PCR with a pMAKD vector backbone made up of 0.05 units/ μ l *Taq* polymerase (Invitrogen, Carlsbad, CA) plus 0.025 units/ μ l PFU polymerase (Stratagene, La Jolla, CA); 10 cycles of 94°C for 15 s, 65°C for 30 s, 68°C for 12 min. After column purification (QIAGEN, Valencia, CA), the PCR product was dissolved in 10 μ l of 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, and 1 μ l was electroporated directly into *E. coli*.





Group II intron-stimulated homologous recombination. To introduce stop codons into the chromosomal *thyA* gene (Fig. 5), the *recA⁺* *E. coli* strain BL21(DE3) was co-transformed with pBR-thyA-Stop 1 or 2 and intron-donor plasmid pACD2-ThyA-495a-RT⁻ and grown overnight at 37°C in LB medium containing chloramphenicol and ampicillin to select for both plasmids. After induction with 100 µM IPTG for 1 h at 37°C, the cells were transferred to fresh LBT (LB medium supplemented with 50 µg/ml thymine), grown for an additional hour for completion of double-strand break repair, and plated on minimal medium containing 50 µg/ml thymine, with and without trimethoprim.

DNA analysis. *E. coli* genomic DNA and plasmids were prepared by using a Genomic DNA isolation kit and Plasmid Maxi kit, respectively (QIAGEN).

For Southern hybridizations, DNAs were digested with restriction enzymes, run in a 0.7% agarose gel, blotted to a nylon membrane, and hybridized as described²⁸ with ³²P-labeled DNA probes prepared by random octamer labeling of PCR products with ³²P-dCTP (RTS RadPrime DNA Labeling System; Invitrogen).

Acknowledgments

This work was supported by NIH grant GM37949. We thank Stephanie Reeves, Shelley Payne, and Rasika Harshey for *Shigella* and *Salmonella* strains, and Marlene Belfort and Keith Derbyshire (Albany) for comments on the manuscript.

Received 15 May 2001; accepted 10 October 2001

- Lambowitz, A.M., Caprara, M.G., Zimmerly, S. & Perlman, P.S. Group I and group II ribozymes as RNPs: clues to the past and guides to the future. In *The RNA world*, Edn. 2. (eds Gesteland, R.F., Cech, T.R. & Atkins, J.F.) 451–484 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; 1999).
- Belfort, M., Derbyshire, V., Parker, M.M., Cousineau, B. & Lambowitz, A.M. Mobile introns: pathways and proteins. In *Mobile DNA*, Edn. 2. (eds Craig, N., Craigie, R., Gellert, M. & Lambowitz, A.M.) (ASM Press, Washington DC; 2002), in press.
- Eskes, R., Yang, J., Lambowitz, A.M. & Perlman, P.S. Mobility of yeast mitochondrial group II introns: engineering a new site specificity and retrohoming via full reverse splicing. *Cell* **88**, 865–874 (1997).
- Cousineau, B. *et al.* Retrohoming of a bacterial group II intron: mobility via complete reverse splicing, independent of homologous DNA recombination. *Cell* **94**, 451–462 (1998).
- Eskes, R. *et al.* Multiple homing pathways used by yeast mitochondrial group II introns. *Mol. Cell. Biol.* **20**, 8432–8446 (2000).
- Cousineau, B., Lawrence, S., Smith, D. & Belfort, M. Retrotransposition of a bacterial group II intron. *Nature* **404**, 1018–1021 (2000).
- Saldanha, R. *et al.* RNA and protein catalysis in group II intron splicing and mobility reactions using purified components. *Biochemistry* **38**, 9069–9083 (1999).
- Guo, H. *et al.* Group II introns designed to insert into therapeutically relevant DNA target sites in human cells. *Science* **289**, 452–457 (2000).
- Guo, H., Zimmerly, S., Perlman, P.S. & Lambowitz, A.M. Group II intron endonucleases use both RNA and protein subunits for recognition of specific sequences in double-stranded DNA. *EMBO J.* **16**, 6835–6848 (1997).
- Mohr, G., Smith, D., Belfort, M. & Lambowitz, A.M. Rules for DNA target-site recognition by a lactococcal group II intron enable retargeting of the intron to specific DNA sequences. *Genes Dev.* **14**, 559–573 (2000).
- Singh, N.N. & Lambowitz, A.M. Interaction of a group II intron ribonucleoprotein endonuclease with its DNA target site investigated by DNA footprinting and modification interference. *J. Mol. Biol.* **309**, 361–386 (2001).
- Belfort, M. & Pedersen-Lane, J. Genetic system for analyzing *Escherichia coli* thymidylate synthase. *J. Bacteriol.* **160**, 371–378 (1984).
- Davanloo, P., Rosenberg, A.H., Dunn, J.J. & Studier, F.W. Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *Proc. Natl. Acad. Sci. USA* **81**, 2035–2039 (1984).
- Jasin, M. Genetic manipulation of genomes with rare-cutting endonucleases. *Trends Genet.* **12**, 224–228 (1996).
- Puchta, H., Dujon, B. & Hohn, B. Two different but related mechanisms are used in plants for repair of genomic double-strand breaks by homologous recombination. *Proc. Natl. Acad. Sci. USA* **93**, 5055–5060 (1996).
- Donoho, G., Jasin, M. & Berg, P. Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol. Cell. Biol.* **18**, 4070–4078 (1998).
- Posfai, G., Kolisnychenko, V., Bereczki, Z. & Blattner, F.R. Markerless gene replacement in *Escherichia coli* stimulated by a double-strand break in the chromosome. *Nucleic Acids Res.* **27**, 4409–4415 (1999).
- Murphy, K.C. Use of bacteriophage λ recombination functions to promote gene replacement in *Escherichia coli*. *J. Bacteriol.* **180**, 2063–2071 (1998).
- Zhang, Y., Buchholz, F., Muyrers, J.P.P. & Stewart, F.A. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.* **20**, 123–128 (1998).
- Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97**, 6640–6645 (2000).
- Murphy, K.C., Campellone, K.G. & Poteete, A.R. PCR-mediated gene replacement in *Escherichia coli*. *Gene* **246**, 321–330 (2000).
- Yu, D. *et al.* An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **97**, 5978–5983 (2000).
- Ellis, H.M., Yu, D., DiTizio, T. & Court, D.L. High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl. Acad. Sci. USA* **98**, 6742–6746 (2001).
- Matsuura, M. *et al.* A bacterial group II intron encoding reverse transcriptase, maturase and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron. *Genes Dev.* **11**, 2910–2924 (1997).
- Payne, S.M., Niesel, D.W., Peixotto, S.S. & Lawlor, K.M. Expression of hydroxamate and phenolate siderophores by *Shigella flexneri*. *J. Bacteriol.* **155**, 949–955 (1983).
- Tsai, S.P., Hartin, R.J. & Ryu, J. Transformation in restriction-deficient *Salmonella typhimurium* LT2. *J. Gen. Microbiol.* **135**, 2561–2567 (1989).
- Hamilton, C.M., Aldea, M., Washburn, B.K., Babbitzke, P. & Kushner, S.R. New method for generating deletions and gene replacements in *Escherichia coli*. *J. Bacteriol.* **171**, 4617–4622 (1989).
- Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular cloning: a laboratory manual*, Edn. 2. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; 1989).