

MINIREVIEW

Barriers to Intron Promiscuity in Bacteria

DAVID R. EDGELL,^{1*} MARLENE BELFORT,² AND DAVID A. SHUB¹

*Department of Biological Sciences and Center for Molecular Genetics, University at Albany—SUNY, Albany, New York 12222,¹
and Molecular Genetics Program, Wadsworth Center, New York State Department of Health and School of Public Health,
University at Albany—SUNY, Albany, New York 12201-2002²*

“The field of bacterial viruses is a fine playground for serious children who ask ambitious questions.”

Max Delbruck

The first bacterial intron, a self-splicing group I intron, was found to interrupt the thymidylate synthase (*td*) gene of the *Escherichia coli* phage T4 (11). The second and third bacterial group I introns were found to interrupt the aerobic (*nrdB*) and anaerobic (*nrdD* [initially named *sunY*]) ribonucleotide reductases of phage T4 (29, 90), and another group I intron was soon discovered in the DNA polymerase gene of SPO1, a *Bacillus* phage (25). From this (admittedly) small sampling of phage genomes, one might have naively expected that group I introns would be abundant in phage or bacterial genomes, especially since subsequent laboratory experiments demonstrated that group I introns could propagate themselves (by a process called homing) throughout populations of intron-minus alleles with near 100% efficiency (5, 68). That a similar homing phenomenon had also been previously demonstrated for a group I intron in the large rRNA gene of yeast mitochondria (34) gave additional support to the notion that group I introns should be able to spread efficiently throughout populations. However, this expected outcome has never been realized in natural phage populations; some phage populations harbor many introns, while other related phage populations are strangely lacking in any introns whatsoever (Table 1). Why do group I introns have an unusual distribution in phage and bacterial genomes, and what potential barriers might exist to prevent their spread?

A similar question might be asked of group II introns in bacteria. Much was made of the initial finding of group II introns in bacteria, as their discovery added fuel to the debate concerning the evolutionary origins of eukaryotic spliceosomal introns (22, 70) which have both structural and functional similarities to group II introns (53, 79, 84). Yet, the number of group II introns in bacteria is small, many of which are inferred only from database matches to reverse transcriptases or maturases encoded within known introns (13, 41, 73, 89), and only two have been shown to splice or be mobile in vivo (48, 49, 55, 80) (Table 2). While group II intron homing is mechanistically distinct from group I intron homing, the principle is similar; group II introns home from intron-containing to intronless alleles. Many elegant biochemical and genetic experiments have unraveled the complexities of bacterial group II intron homing (14, 48, 49), and based on these results, there seems no a priori reason why group II introns should not be able to

spread efficiently through populations of intron-minus alleles. The paucity of bacterial group II introns becomes even more perplexing given the recent demonstration of group II intron transposition to novel chromosomal sites (15). That group II introns are abundant in mitochondrial and chloroplast genomes (52) and present in bacterial genomes but at lower levels (and absent in phages) only adds to the mystery surrounding the lack of group II introns in bacteria, as mitochondria and chloroplasts are typically prokaryotic in genome organization and ultimately are derived from two distinct bacterial lineages. Do similar barriers that prevent group I introns from spreading throughout bacterial populations also apply to preventing the spread of group II introns in bacteria?

WHAT IS INTRON HOMING?

Homing is the process by which group I and II introns spread through a population of homologous (or cognate) intronless alleles (17) (Fig. 1). Homing is completely dependent on and initiated by intron-encoded proteins, and for group I and group II introns, completion of the homing process relies on host-encoded proteins (for a review, see reference 42). Group I intron homing is initiated by a DNA endonuclease (usually called a homing endonuclease) encoded within the intron itself. Four families of endonucleases were initially identified on the basis of conserved amino acid motifs and structural similarities; they are the GIY-YIG (9, 51), LAGLIDAG (31, 50), His-Cys box (60), and H-N-H (28, 83) families (for reviews, see references 3 and 4). However, structural similarities in active-site architecture have recently been found in representative H-N-H and His-Cys box endonuclease crystal structures, suggesting that these two families could possibly be merged into one (39). Homing endonucleases of all families recognize lengthy DNA sequences (14 to 40 bp) that span the analogous intron insertion site in the intronless allele (4) and mostly introduce a double-strand break (DSB). Since the recognition sequence of the intron-encoded endonuclease is interrupted by the intron, intron-containing alleles are immune to cleavage by their own endonucleases.

In group I intron homing, host- and phage-encoded DNA repair proteins use the intron-containing allele as a template to repair the DSB in the intronless allele, thus copying the intron and its endonuclease gene into the intron-minus allele (Fig. 1). In the *E. coli* bacteriophage T4, two repair pathways, DSB repair (DSBR) and synthesis-dependent strand annealing (SDSA) are implicated in repair of the DSB during intron acquisition (57). These repair pathways differ in that resolution of the Holliday junction in DSBR results in both crossover and noncrossover products, whereas no Holliday junctions are found and no crossover products are generated in the SDSA pathway.

* Corresponding author. Mailing address: Molecular Genetics Program, Wadsworth Center, New York State Department of Health and School of Public Health, University at Albany—SUNY, Albany, NY 12201-2002. E-mail: buddy@csc.albany.edu.

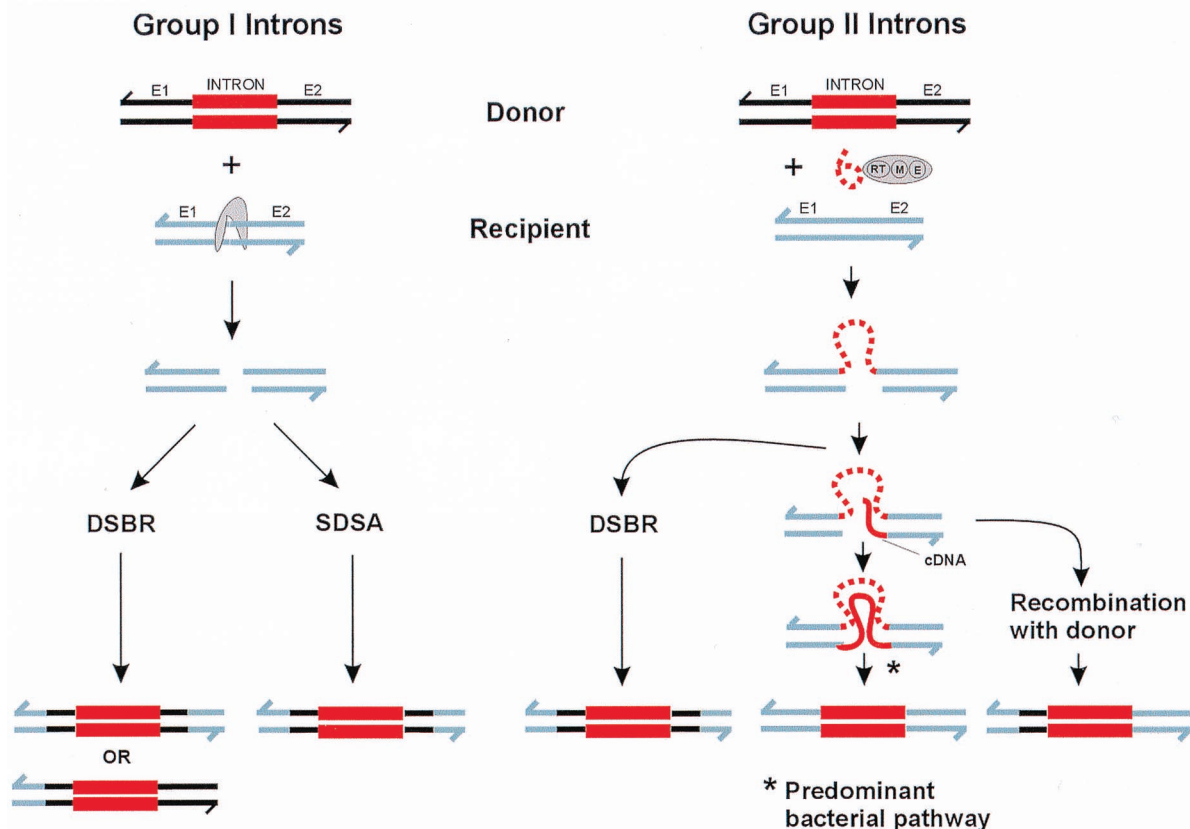


FIG. 1. Schematic of group I and II intron homing pathways. Solid black lines represent upstream (E1) and downstream (E2) exons of donor (intron-containing) alleles, and blue-gray lines indicate recipient (intronless) alleles. Group I or II introns are depicted by red boxes. Broken red lines indicate intron RNA, while solid red lines indicate cDNA synthesized by the reverse transcriptase (RT) function of group II intron proteins. The gray tooth-shaped structure represents a group I intron endonuclease. Group II intron proteins possess two functions in addition to RT, namely, maturase (M) and endonuclease (E) activities.

For group II introns, remarkably, the intron RNA forms a complex with the intron-encoded protein to initiate homing by recognizing a specific DNA sequence (Fig. 1). The intron RNA is an essential catalytic component of this process (49, 76, 93). Because group II intron homing can proceed through a reverse-transcribed copy of the intron RNA, the process has been termed retrohoming to distinguish it from endonuclease-mediated homing of group I introns, which proceeds strictly through DNA intermediates (16). And, unlike group I intron homing where the intron-encoded endonuclease is sufficient to promote homing into intron-minus alleles, group II intron homing relies on two functions of the intron-encoded protein in addition to the endonuclease function; maturase, which promotes intron splicing, and reverse transcriptase, which generates intermediates of the retrohoming process (reviewed in reference 43). Whereas in bacteria, group II intron proteins are encoded by self-contained open reading frames (ORFs) with independent promoters (91), in yeast mitochondria, they are produced as fusion proteins with the upstream exon (75). Subsequent processing of this polyprotein yields a product that initially functions as a maturase to promote stabilization of intron structures necessary for splicing. Splicing of the intron from the pre-mRNA releases an intron lariat that forms a RNP with the intron-encoded protein (43). This complex binds to the target sequence in the intron-minus alleles where the intron RNA cleaves the sense strand of the DNA at the intron insertion site by a reverse splicing reaction, and the intron becomes covalently attached to the intron-minus allele. Cleav-

age of the antisense strand is carried out by the endonuclease function of the intron-encoded protein, and the 3' end serves as the primer for cDNA synthesis by the reverse transcriptase function of the intron-encoded protein using the integrated intron as the template (92).

It is at this point in the homing process where bacterial and yeast mitochondrial group II homing can potentially differ (Fig. 1). Retrohoming of the *Lactococcus lactis* Ll.LtrB intron proceeds independent of homologous recombination (14), whereas the yeast introns require recombination (20). In the bacterial pathway, the intron RNA undergoes a reverse splicing reaction and serves as a template for complete cDNA synthesis (14). In contrast, and depending on the yeast mitochondrial group II intron, cDNA synthesis may not proceed to the end of the intron, and completion of the retrohoming process is thus dependent on the cDNA invading the intron donor allele and on homologous recombination between the donor and recipient alleles (20). Recombination results in co-conversion of flanking exon sequences, a result not observed in the bacterial retrohoming pathway.

WHERE ARE GROUP I AND II INTRONS FOUND?

Although group I and II introns are found in a diverse phylogenetic sampling of bacteria and phages, both exhibit an unusual genomic distribution (Tables 1 and 2). For group I introns, three interesting themes emerge. First, group I introns do not interrupt protein-coding genes in bacterial chromo-

TABLE 1. Distribution of group I introns in bacteria and bacteriophages

Bacterium or bacteriophage	Species	Gene or enzyme	Endonuclease ^a	Splicing ^b	Reference(s)
Bacteria					
Gram-negative	<i>A. tumefaciens</i>	tRNA ^{Arg} _{CCU}	None	Yes	69
	<i>Azoarcus</i> sp. strain BH72	tRNA ^{Ile} _{CAU}	None	Yes	69
Chlamydial	<i>S. negevensis</i> Z ^T	23S rRNA	LAGLIDAG	No	21
Cyanobacterial	Many	tRNA ^{Leu} _{UAA}	None	Yes	40, 64, 74, 87
	Many	tRNA ^{fMet}	None ^c	Yes	6, 64
Bacteriophages					
Gram-negative bacteria	<i>E. coli</i> /T4, T6, RB3, LZ2, Tu1A	Thymidylate synthase (<i>td</i>)	I-TevI (GIY-YIG)	Yes	11, 29, 67
	<i>E. coli</i> /T4	Anaerobic ribonucleotide reductase (<i>nrdD</i>)	I-TevII (GIY-YIG)	Yes	29, 67
	<i>E. coli</i> /T4, RB3	Aerobic ribonucleotide reductase (<i>nrdB</i>)	Fragment	Yes	67
Gram-positive bacteria	<i>B. subtilis</i> /β22	<i>td</i>	Fragment	Yes	1
	<i>B. subtilis</i> /SPO1	DNA polymerase	I-HmuI (H-N-H)	Yes	25, 26, 27
	<i>B. subtilis</i> /SP82	DNA polymerase	I-HmuII (H-N-H)	Yes	26, 27
	<i>B. subtilis</i> /φe	DNA polymerase	I-HmuIII (H-N-H)	Yes	26
	<i>B. subtilis</i> /SPβ prophage	Ribonucleotide reductase (<i>nrdF</i>)	<i>yosQ</i> (H-N-H)	Yes	45
	<i>S. thermophilus</i> /Sb3	Lysin	H-N-H	Yes	24
	<i>L. lactis</i> /r1t	ORF40	H-N-H	Yes	61
	<i>Lactococcus delbrueckii</i> /LL-H	Terminase	ORF168 (H-N-H)	Yes	54
	<i>Staphylococcus aureus</i> /Twort	ORF142 introns 1, 2, 3	None	Yes	44

^a When known, endonuclease names are given (e.g., I-TevI). "Fragment" refers to a truncated ORF, and "None" means no ORF was present in the intron. Endonucleases are also identified by which family they belong to (GIY-YIG, H-N-H, His-Cys Box, or LAGLIDAG).

^b For the majority of group I introns, evidence for splicing is based on in vitro assays. Only a few introns have been shown to be mobile in vivo.

^c One tRNA^{fMet} gene, that of *Synechocystis* sp. strain 6803, contains a novel endonuclease, I-SspI (R. Bonocora and D. A. Shub, submitted for publication).

somes, they interrupt only tRNA genes (40, 64, 65, 74, 87; information found at the Comparative RNA website of The Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin [http://www.rna.icmb.utexas.edu]). There are 74 known bacterial chromosomal group I introns, and 73 interrupt tRNA genes (information found at the Comparative RNA website). However, most of these introns are very similar, being isolated from closely related species or different strains of the same species of bacterium. In total,

four different tRNAs are interrupted by introns, tRNA^{Leu}_{UAA}, tRNA^{Arg}_{CCU}, tRNA^{fMet}, and tRNA^{Ile}_{CAU}. There is only one example of an interrupted bacterial rRNA gene, that of the 23S ribosomal DNA (rDNA) of *Simkania negevensis* Z^T (21). Although the interrupting sequence can be folded into a secondary structure very similar to group I introns of the IB4 subgroup, typical of organellar rRNA introns and atypical of bacterial tRNA introns, there is no evidence for splicing in vivo as assayed by reverse transcription-PCR. The single example of

TABLE 2. Distribution of group II introns in bacteria

Bacterium	Species ^a	Intron/insertion site ^b	ORF	Splicing ^c	Reference(s)
Gram-negative	<i>E. coli</i> ECOR 43, 17, 21, 31, 67	IntA/H-repeat element	Yes	ND	23
	<i>E. coli</i> ECOR 44, 47	IntB/H-repeat element	Yes	ND	23
	<i>E. coli</i> ECOR 9, 52, 60, 17, 20, 23, 24	IntC/plasmid-related ORF	Yes	ND	23
	<i>E. coli</i> ECOR 9, 38	IntD/IS3411	Yes	ND	23
	<i>E. coli</i>	ND/pR471a inferred—retron?	Yes	ND	41
	<i>Azotobacter vinelandii</i> UWR	AvUWRx.1/ND	??	ND	22
	<i>Sphingomonas aromaticivorans</i> F199	ND/ND	Yes	ND	73
Cyanobacterial	<i>Pseudomonas alcaligenes</i> NCIB 9867	Xln6/pRP4	Yes	Inferred	89
	<i>Calothrix</i> sp. strain PCC7601	Cal.x1/ND	Yes	Yes	22
	<i>Calothrix</i> sp. strain PCC7101	Cal.x2/ssDNA binding protein ^d	Yes	ND	22
Gram-positive	<i>L. lactis</i> ML3	LI.LtrB/relaxase gene, pRS01	Yes	Yes	55, 56
	<i>L. lactis</i> 712	IntL/conjugation gene <i>mobA</i>	Yes	Yes	80
	<i>S. meliloti</i> GR4	<i>RmInt1/ISRm2011-2</i> , pRmeGR4b	Yes	Yes	48
	<i>Streptococcus pneumoniae</i>	ND/ND	Yes	ND	13

^a Some group II introns were identified from various members of the *E. coli* ECOR collection, as listed by their strain number.

^b Insertion sites of some group II introns were not identified, indicated by ND.

^c There is no evidence for splicing for the majority of putatively identified group II introns, indicated by ND.

^d ssDNA, single-stranded DNA.

a chromosomally encoded group I intron in a protein-coding gene is in the ribonucleotide reductase gene (*nrdF*) of SP β , an integrated prophage of *Bacillus subtilis* (45).

Second, the converse is true for phage-encoded group I introns; these introns are always found interrupting protein-coding genes, never structural RNA genes (some phages carry tRNA genes). Furthermore, most group I phage introns interrupt highly conserved protein-coding genes, many of which are involved in some aspect of DNA metabolism. This discovery prompted speculation that the function of phage introns might be the regulation of DNA metabolism, thus accounting for their retention in size-constrained phage genomes (25, 81; see below).

Third, homing endonucleases are typically encoded within phage group I introns, not in chromosomal group I introns, with two exceptions. The first exception is the tRNA^{fMet} intron of *Synechocystis* that contains an ORF that potentially codes for a small basic protein not similar to any known endonuclease (6), and the second is a LAGLIDAG endonuclease ORF within the *S. negevensis* Z^T 23S rDNA insertion (21). However, in spite of the tendency for phage introns to encode endonucleases, not all phage introns do, and some appear to have remnants of ORFs, evidenced by the phage T4 *nrdB* intron (19) and the *Bacillus* phage β 22 thymidylate synthase intron (1). Perhaps the best example of endonuclease-lacking phage introns is found in the *Staphylococcus* phage Twort; three group I introns interrupt a late-transcribed gene (ORF142) of unknown function, and none of them code for endonucleases (44). Generally then, phage introns that interrupt highly conserved protein-coding genes code for homing endonucleases, whereas bacterial chromosome-encoded group I introns interrupt tRNA genes and do not code for endonucleases.

Although there are fewer known examples of group II introns than of group I introns in bacteria, one trend is immediately clear: most group II introns are inserted in, or associated with, other mobile genetic elements (23, 32, 48, 56, 59, 80) (Table 2). The best studied bacterial group II intron, the Ll.LtrB intron of *L. lactis* ML3, is inserted in the conjugative relaxase gene essential for transfer of the plasmid pRS01 (56). The same intron, but designated *IntL*, was discovered independently in a chromosomally encoded gene, *mobA*, required for conjugation in *L. lactis* 712 (80). Another group II intron, RmInt1, is inserted within an insertion sequence of the IS630-Tc1/IS3 retroposon superfamily that is found on a plasmid, pRmeGR4b, of *Sinorhizobium meliloti* (48). Four group II introns have been identified in isolates of the *E. coli* ECOR collection, and all of them are associated with some form of mobile DNA (23, 36). Two introns from distinct ECOR isolates are inserted at different positions within the mobile Hinc repeat of *Rhs* elements. A third *E. coli* intron, found in two different ECOR strains, is inserted within a sequence that is very similar to IS3411, and the fourth intron is inserted in an ORF of unknown function but is related to an ORF encoded by T-DNA of *Agrobacterium tumefaciens*.

DISTRIBUTION OF GROUP I INTRONS IN PHAGE POPULATIONS

It is perhaps ironic that self-splicing group I introns were discovered as a property of T-even phages of gram-negative bacteria, for gram-negative bacterial phages are where group I introns are least abundant (18). Apart from the *td*, *nrdD*, and *nrdB* introns in T-even phages, no other group I introns have been discovered in any gram-negative bacterial phage. Conversely, group I introns are more abundant in phages of gram-positive bacteria (1, 24, 25, 26, 45, 54, 61), exemplified by the

Staphylococcus phage Twort, which possesses at least five introns (44). While it is tempting to generalize about group I intron distribution in phages, it is worth noting that only two detailed studies have been performed; one concerned with distribution of the three T4 introns in related T-even phages (18) and the other concerned with the distribution of a lysin intron in *Streptococcus thermophilus* phages (24).

Phage T4 is unusual with respect to other T-even phages in that it possesses three group I introns. Two of the three, the *td* and *nrdD* introns, are mobile, while the third intron, *nrdB*, is not (68). T-even phage RB3 also harbors a *nrdB* intron very similar to the T4 intron, but the RB3 intron is almost twice as long because it possesses a complete ORF coding for a functional endonuclease (which T4 *nrdB* does not [19]). The T4 *nrdB* intron thus appears to carry a deletion of the ORF relative to the RB3 *nrdB* intron, but the endonuclease was presumably intact and functional at some point. Other T-even phages are also surprisingly bereft of introns; for instance, phage T6 has only one intron, *td*, and T2 has no introns (67). (One T2 strain, T2W, harbors both the *td* and *nrdD* introns. However, experimental evidence suggests that these introns were inherited during laboratory propagation of T2 and T4.)

In an effort to better understand the distribution of introns in natural T-even phage populations, Eddy screened 32 phages, a number of which were novel isolates of mammalian fecal matter from the Denver Zoo or from clarified sewage samples of city water treatment plants (18). The results are surprising. Using a combination of GTP labeling of RNA, primer extension sequencing, and PCR fingerprinting, Eddy was able to detect a total of eight intron occurrences in all phages examined. In phage isolates from which PCR products were obtained, the *nrdD* and *nrdB* introns appear to be very rare, present in only 1 of 30 and 2 of 31 phages, respectively. However, the *td* intron is more widespread, present in 5 of 22 phages from which a thymidylate synthase gene could be amplified. Also, Eddy found that every phage isolate except one could infect at least four *E. coli* strains differing in cell surface receptors. Thus, it seems unlikely that the paucity of introns in T-even phages is due to lack of opportunity of intron-carrying phages to infect similar host cell populations as intronless phages.

The second detailed study of group I intron distribution reaches a somewhat different conclusion than that of the above survey (24). *S. thermophilus* is an economically important bacterium for the dairy industry, and as such, its phages have been subject to much study. Five independent phage isolates have been completely sequenced, allowing a detailed PCR survey of related phages for the presence of introns. Foley and coworkers initially found a self-splicing group I intron, complete with an H-N-H endonuclease ORF, interrupting the lysin gene of phage Sf1 (24). When they expanded their PCR survey to encompass 61 phages isolated from geographically and ecologically distinct locales, they found that 31 of the phages contained the intron, four of which have deletions within the H-N-H ORF. Thus, it appears that intron spread in this phage population has been much more successful than intron spread in populations of T-even phages.

T-EVEN INTRONS: RECENT ARRIVALS OR LONG-TERM RESIDENTS?

Of the distribution of introns in T-even phages, Eddy (18) commented that "Although variable intron occurrence has been said to imply that the introns are spread through the population by virtue of their mobility, it actually seems somewhat puzzling that these highly infectious introns do not occur

in most or all T-even isolates." What then accounts for the sporadic and limited distribution of introns in the T-even phage family? One argument is that the introns have only recently arrived in T-even phages and have not yet had sufficient time or opportunity to spread throughout other T-even phages. However, we present the following points supporting an alternative view that the T4 introns have been present for a significant period of time in the T-even population (2).

First, expression of the three T4 intron-encoded ORFs is tightly regulated (30). All three intron-containing genes are expressed as early or middle transcripts from promoters upstream of the respective genes, but translation of the intron endonuclease gene from the pre-mRNA is prevented by the presence of secondary structures that occlude ribosome binding sites. As intron endonucleases are expressed from late promoters found within the introns, these RNA transcripts do not include secondary structures to occlude the endonuclease gene ribosome binding site, thus facilitating translation of the endonuclease. This mode of regulation must be the result of specific adaptation in T4 or in another T-even-like phage, since the highly unusual T4 late promoter sequences are present within the introns themselves. Second, the intron-encoded endonucleases *I-TevI*, *I-TevII*, and *I-TevIII* appear to have adapted to the T-even environment as codon usage is optimized for T4 (82). Third, both *I-TevI* and *I-TevII*, the *nrdD* intron endonuclease, make predominantly minor groove contacts with their DNA recognition sequences (7, 47). This mode of DNA contact is likely an adaptation to modified T-even phage DNA, which in T4 contains bulky α - and β -glucosylated hydroxymethyl cytosine residues in the major groove (46). These adaptations are counter to the introns and endonucleases being recent arrivals in the T-even phage population.

WHY ARE PHAGE INTRONS IN THE GENES THAT THEY ARE IN?

The startling discovery that phage T4 group I introns were inserted within genes of nucleotide metabolism (thymidylate synthase [*td*], aerobic [*nrdB*], and anaerobic [*nrdD*] ribonucleotide reductases) led to the proposal that the function of these introns was to regulate levels of nucleotide precursors for DNA synthesis through regulation of splicing (81). This proposal was strengthened by the discovery of another group I intron in a gene involved in DNA metabolism, the DNA polymerase of *Bacillus* phage SPO1 (25). Thus, a rationale was provided for the retention of seemingly optional introns in size-constrained genomes subject to pressures of genome streamlining for rapid replication. However, this presumed function for the T4 introns began to lose attractiveness with the discovery that close relatives of T4 lacked introns in these genes (67), with the (unpublished) experimental observation that T4 introns could be deleted without detectable detriment to the phage (18), and with the finding that some phage introns do not interrupt genes of nucleotide metabolism (44, 54, 61).

Why then do phage introns tend to be inserted within highly conserved genes involved in some aspect of DNA metabolism? One possible reason relates to the relatively high specificity of group I intron-encoded endonucleases. In general, homing endonucleases recognize relatively long stretches of DNA (14 to 40 bp) and, depending on the individual endonuclease, tolerate various numbers of nucleotide substitutions within their respective recognition sequences (35). Yet, homing endonucleases in phage populations are presented with a limited number of potential recognition sequences because phages have relatively small genomes. Small genome size also constrains the phage with respect to number of genes it can ultimately en-

code, so phages tend to carry genes coding essential functions: proteins involved in replication, transcription, and morphogenesis. Thus, for phage-encoded homing endonucleases, one possible strategy to maximize spread to related phages comes from using conserved nucleotide sequences, which are likely to lie within genes that function in replication and transcription, as recognition sites.

Furthermore, a specific prediction would be that the recognition sites of homing endonucleases are nucleotide sequences corresponding to important functional domains within those proteins, as these sequences are more likely to be conserved between phages. For instance, the recognition sequence of *I-TevI*, the phage T4 thymidylate synthase intron endonuclease, encompasses 37 bp (7, 8) that includes codons for the highly conserved Arg218, Ser219, Asp221, and Asn229 (*E. coli* numbering), all of which are functionally critical (10). Thymidylate synthase catalyzes the reductive methylation of dUMP by 5,10-methylene-5,6,7,8-tetrahydrofolate to produce dTMP and 7,8-dihydrofolate. In the crystal structure of the *E. coli* enzyme complexed with dUMP, Arg218 and Ser219 form one hydrogen bond each with the phosphate of dUMP. Ser219 and Asp221 are thought to be important for the enzyme's reduced affinity for ribose nucleotides because these residues would sterically hinder binding of nucleotides containing a 2'-OH group. In addition, Asp221 and Asn229 each form hydrogen bonds with the pyrimidine ring of dUMP. That other homing endonucleases also have recognition sequences that encompass functionally critical residues can be easily verified as additional structures and mutational data become available.

Intron insertion into essential phage genes also imposes negative selection on the phage to remove the intron, since inexact deletion of an intron from an essential gene may potentially affect codons corresponding to amino acid residues critical for function. Ironically, precise deletion of the intron from an individual phage within a population would provide positive selection for retention of the endonuclease ORF, because the homing site would be regenerated and become a substrate for endonucleases of other phage within the same population. Of course, no hypothesis is bulletproof, and there are already examples of endonuclease-containing introns interrupting genes not involved in replication or transcription (the lysin gene of *S. thermophilus* phages [24] and the terminase gene of *Lactobacillus* phage LL-H [54]). It is possible, however, that these endonucleases are also targeting conserved nucleotide sequences within their respective genes.

HOW DO BACTERIAL INTRONS GET AROUND?

Intron homing is an efficient but confining process that results in the spread and potential fixation of a group I or II intron in a population of intronless alleles. Nevertheless, homing does not promote intron dissemination to nonallelic sites. Intron transposition, however, can do this.

Intron transposition to nonallelic sites has been inferred from well-documented studies for fungal mitochondrial group II introns (58, 77). Group II intron transposition in bacteria was also inferred from studies on the distribution and insertion sites of an intron in strains of *S. meliloti* (48). Recently, Cousineau and coworkers experimentally demonstrated retrotransposition for the first time, showing movement of the *L. lactis* Ll.LtrB group II intron to novel sites in the chromosome (15). Unlike retrohoming, retrotransposition can occur independently of the endonuclease function of the intron-encoded protein but is dependent on host recombinase functions. Furthermore, retrotransposition events resulted in introns with various degrees of splicing function, as judged from primer

extension analysis of *L. lactis* RNA isolated from strains harboring novel intron insertions. Given these results, it is curious and surprising that group II introns are not more widely distributed throughout bacterial or phage genomes.

What evidence, if any, exists for group I intron transposition in bacteria? In vitro and in vivo experimental evidence suggests that reverse splicing of group I intron RNA into nonallelic RNA is one possible mechanism of intron transposition (85). Because as little as 4 to 6 bp of sequence complementarity is required between the intron and target RNA to initiate a reverse splicing event, many potential sites exist within any population of RNA molecules. Indeed, Roman and Woodson (71) have demonstrated reverse splicing in vivo of the *Tetrahymena* large-subunit (LSU) RNA group I intron into *E. coli* 23S RNA at 11 novel sites. The abundance of structural RNA transcripts (rRNAs and tRNAs) has led to the suggestion that such RNAs would be preferential targets for reverse splicing and transposition, providing a possible explanation for the prevalence of group I introns in bacterial tRNA genes (3). Nevertheless, the limiting step in reverse-splicing-mediated group I intron transposition must surely be reverse transcription of the intron and target RNA and subsequent recombination, given that the requisite reverse transcriptase enzyme activity is not prevalent in many bacterial cells.

The relaxed sequence specificity of some intron-encoded endonucleases may also provide a means by which group I introns could invade new alleles (7, 8), as DSBs might result from the recognition and cleavage of degenerate versions of homing sites by intron endonucleases. Whether the same pathways that repair DSBs in allelic homing events could repair DSBs generated at nonallelic sites is determined by the amount of homology between flanking exon sequences. Detailed studies on the homology requirements for T4 *td* intron homing suggest that efficient homing is supported by exon homologies of 50 bp or greater. However, homing events were still observed with limiting exon homology (10 bp) or when homology is confined to only one exon in T4 phages carrying a mutation in the 3'-5' exonuclease *DexA*, which functions to degrade 3' tails of DSBs (33, 66). Illegitimate recombination may be involved in the repair of DSBs with limiting exon homology and thus provide a potential pathway for group I intron invasion at nonallelic sites.

HOW EFFICIENT IS INTRON TRANSPOSITION?

One factor that must influence intron dissemination in bacteria is the frequency and efficiency of group I and II intron transposition versus the frequency and efficiency of homing. Homing is an extremely efficient process, driven in part by the high affinity of group I intron endonucleases or group II intron RNP complexes for their respective recognition sequences, resulting in near 100% conversion of recipient intron-minus alleles to intron-plus alleles. In contrast, transposition is an infrequent process due, in part, to the reduced affinity of intron-encoded proteins for variants of their recognition sequences and to heterology between intron-containing and intronless alleles. Furthermore, for phage-encoded endonucleases, whose expression is temporally restricted (30) and of low abundance, opportunities for transposition are probably rare.

Intron transposition, whether mediated by reverse splicing or by an endonuclease, places the intron in a nonnative environment, with potential effects on intron splicing. For instance, although the *Tetrahymena* LSU RNA group I intron can reverse splice into novel sites in *E. coli* 23S RNA, the forward-splicing efficiency is 10-fold lower than from its natural site (72). Likewise, group II intron transposition to novel locations

in *L. lactis* resulted in forward-splicing rates lower than those observed from the intron's natural splice junction (15). These results suggest that RNA sequence and structure can potentially influence sites at which transposition events result in splicing-competent introns.

WHAT MIGHT SELECT AGAINST INTRON INSERTION INTO CHROMOSOMAL GENES?

Without exception, group I introns found in bacterial chromosomes are never inserted in protein-coding genes (Table 1). Why is this so? One possibility immediately comes to mind, which has been suggested by others (11, 62, 78, 81, 86). In bacteria, unlike eukaryotic nuclei, transcription (hence splicing) and translation are coupled. For introns that encode endonucleases and thus exceed 1,000 bp in length, the ribosome is positioned over the 5' splice site of the intron RNA before RNA polymerase has reached the 3' splice site. Movement of the ribosome into the intron may prevent formation of helices that stabilize the catalytic core of the intron and potentially delay splicing of the intron (63). Group I introns inserted in phage protein-coding genes have overcome this problem by the presence of a stop codon almost immediately after the start of the intron sequence at the 5' splice boundary (or no further than structural element P2) (11, 82). Presumably, this stop codon acts to release the ribosome from the RNA, thus freeing the intron RNA to fold into a catalytically active structure. Ironically, the ribosome may also aid in the folding of the intron core by preventing spurious secondary or tertiary interactions between sequences in the 5' exon and intron RNA (78). Introns inserted in tRNA or rRNA genes do not face this problem of coupled transcription and translation, which may account for the prevalence of group I introns in bacterial tRNA genes.

Nikolcheva and Woodson have shown that group I introns are potentially toxic to bacteria because of possible inhibition of ribosome formation by the intron RNA (62). Using a heterologous system in which a splicing-defective *Tetrahymena* rRNA group I intron was inserted into the analogous position of the *E. coli* 23S RNA, they demonstrated that unspliced 23S RNA accumulated in 50S ribosomes, and these complexes failed to assemble complete 70S ribosomes. In addition, using intron constructs competent for splicing, spliced intron RNA remained noncovalently attached to the 50S ribosome particle, probably through interaction of the intron internal guide sequence with the splice junction in 23S RNA. This association was correlated with slow growth of bacterial strains harboring the group I intron. Thus, there may be strong selection against group I intron insertion into rRNA genes in bacteria.

DOES NONSPECIFIC ENDONUCLEASE ACTIVITY INFLUENCE INTRON DISTRIBUTION?

Ironically, the very same relaxed sequence specificity of group I intron endonucleases that affords them the ability to bind and cleave variants of their recognition sequences may limit genomic retention because of the potentially toxic effects of extraneous DSBs in the host chromosome. One of the more curious observations concerning bacterial group I introns is that, with one experimentally demonstrated exception, endonuclease-containing introns are not found in bacterial chromosomes (Table 1). Yet, the increased genome size of bacteria relative to that of phages could be of benefit only to endonuclease-containing introns, since the number of potential targets, and thus opportunities for transposition, is increased. DSBs generated at ectopic sites on the bacterial chromosome

by endonucleases binding to and cleaving degenerate versions of their recognition sequences must be repaired for the cell to survive, and repair is initiated by homologous sequences on daughter chromosomes. However, laboratory wild-type strains of *E. coli* lack the proper enzyme activities to efficiently repair DSBs (12). Furthermore, when grown in minimal medium, potential repair of any DSB in *E. coli* is further hindered by the lack of multiple chromosomes (38). Perhaps it is not surprising then that endonuclease-containing introns have been most successful in colonizing smaller, multicopy genomes, such as those found in phages, mitochondria, or plastids.

AMBITIOUS QUESTIONS, SOME ANSWERS, MORE QUESTIONS

One problem in attempting to provide reasons for the current genomic and phylogenetic distribution of group I and II introns in bacteria is that there is no expectation as to what intron distribution should be. Efficient homing of group I and II introns in laboratory experiments does not necessarily imply efficient homing in natural populations. Little is known about the dynamics of natural populations of phages (or bacteria) or the frequency of coinfection of the same host cell, factors that surely affect opportunities for intron homing and/or transposition. Studies of intron dynamics at the population level, such as those which highlight surprising differences in group I intron distribution of gram-negative and gram-positive bacterial phages, only emphasize difficulties in applying lessons learned about intron homing from laboratory experiments to natural phage populations. Our view of intron distribution in bacteria is also influenced by laboratory experiments that examine homing events over a limited number of phage or bacterial generations and then the experiments are discarded. Homing events that occur in natural populations must become fixed in bacteria or phages, yet nothing is known about the selective pressures that influence this process.

It is probably correct to state that there is no single reason that can adequately explain the current distribution of group I and II introns in bacteria. Undoubtedly, the number of bona fide group I and II introns in bacteria and phages will increase as more and more prokaryotic genomes are sequenced, as RNA folding and search algorithms become more sophisticated, and as already putatively identified introns are confirmed experimentally. Yet basic questions and observations concerning intron distribution in natural bacterial and phage populations will still apply, and one obvious and exciting future avenue of research is to experimentally address some of the factors that influence intron dissemination in bacteria and their phages.

ACKNOWLEDGMENTS

We thank Archana Belle, Rick Bonocora, Benoit Cousineau, Vicky Derbyshire, Cheryl Eifert, Michelle Gilson, Markus Landthaler, and Niles Lehman for discussion and comments on the manuscript.

D.R.E. was supported by a postdoctoral fellowship from the Medical Research Council of Canada, M.B. was supported by NIH grants GM39422 and GM44844, and D.A.S. was supported by NIH grant GM37746.

REFERENCES

1. **Bechhofer, D. H., K. K. Hue, and D. A. Shub.** 1994. An intron in the thymidylate synthase gene of *Bacillus* bacteriophage β 22: evidence for independent evolution of a gene, its group I intron, and the intron open reading frame. *Proc. Natl. Acad. Sci. USA* **91**:11669–11673.
2. **Belfort, M.** 1989. Bacteriophage introns: parasites within parasites? *Trends Genet.* **5**:209–213.
3. **Belfort, M., and P. S. Perlman.** 1995. Mechanisms of intron mobility. *J. Biol. Chem.* **270**:30237–30240.
4. **Belfort, M., and R. J. Roberts.** 1997. Homing endonucleases: keeping the house in order. *Nucleic Acids Res.* **25**:3379–3388.
5. **Bell-Pedersen, D., S. Quirk, J. Clyman, and M. Belfort.** 1990. Intron mobility in phage T4 is dependent upon a distinctive class of endonucleases and independent of DNA sequences encoding the intron core: mechanistic and evolutionary implications. *Nucleic Acids Res.* **18**:3763–3770.
6. **Biniszkiwicz, D., E. Cesnaviciene, and D. A. Shub.** 1994. Self-splicing group I intron in cyanobacterial initiator methionine tRNA: evidence for lateral transfer of introns in bacteria. *EMBO J.* **13**:4629–4635.
7. **Bryk, M., S. M. Quirk, J. E. Mueller, N. Loizos, C. Lawrence, and M. Belfort.** 1993. The *td* intron endonuclease I-*TeV*I makes extensive sequence-tolerant contacts across the minor groove of its DNA target. *EMBO J.* **12**:2141–2149.
8. **Bryk, M., M. Belisle, J. E. Mueller, and M. Belfort.** 1995. Selection of a remote cleavage site by I-*TeV*I, the *td* intron-encoded endonuclease. *J. Mol. Biol.* **247**:197–210.
9. **Burger, G., and S. Werner.** 1985. The mitochondrial URF1 gene in *Neurospora crassa* has an intron that contains a novel type of URF. *J. Mol. Biol.* **186**:231–242.
10. **Carreras, C. W., and D. V. Santi.** 1995. The catalytic mechanism and structure of thymidylate synthase. *Annu. Rev. Biochem.* **64**:721–762.
11. **Chu, F. K., G. F. Maley, F. Maley, and M. Belfort.** 1984. Intervening sequence in the thymidylate synthase gene of bacteriophage T4. *Proc. Natl. Acad. Sci. USA* **81**:3049–3053.
12. **Clyman, J., and M. Belfort.** 1992. *trans* and *cis* requirements for intron mobility in a prokaryotic system. *Genes Dev.* **6**:1269–1279.
13. **Coffey, T. J., M. C. Enright, M. Daniels, J. K. Morona, R. Morona, W. Hryniewicz, J. C. Paton, and B. G. Spratt.** 1998. Recombinational exchanges at the capsular polysaccharide biosynthetic locus lead to frequent serotype changes among natural isolates of *Streptococcus pneumoniae*. *Mol. Microbiol.* **27**:73–83.
14. **Cousineau, B., D. Smith, S. Lawrence-Cavanagh, J. E. Mueller, J. Yang, D. Mills, D. Manias, G. Dunny, A. M. Lambowitz, and M. Belfort.** 1998. Retrohoming of a bacterial group II intron: mobility via complete reverse splicing, independent of homologous DNA recombination. *Cell* **94**:451–462.
15. **Cousineau, B., S. Lawrence, D. Smith, and M. Belfort.** 2000. Retrotransposition of a bacterial group II intron. *Nature* **404**:1018–1021.
16. **Curcio, M. J., and M. Belfort.** 1996. Retrohoming: cDNA-mediated mobility of group II introns requires a catalytic RNA. *Cell* **84**:9–12.
17. **Dujon, B., M. Belfort, R. A. Butow, C. Jacq, C. Lemieux, P. S. Perlman, and V. M. Vogt.** 1989. Mobile introns: definition of terms and recommended nomenclature. *Gene* **82**:115–118.
18. **Eddy, S. R.** 1992. Introns in the T-even bacteriophages. Ph.D. thesis. University of Colorado, Boulder.
19. **Eddy, S. R., and L. Gold.** 1991. The phage T4 *nrdB* intron: a deletion mutant of a version found in the wild. *Genes Dev.* **5**:1032–1041.
20. **Eskes, R., J. Yang, A. M. Lambowitz, and P. S. Perlman.** 1997. Mobility of yeast mitochondrial group II introns: engineering a new site specificity and retrohoming via full reverse splicing. *Cell* **88**:865–874.
21. **Everett, K. D., S. Kahane, R. M. Bush, and M. G. Friedman.** 1999. An unspliced group I intron in the 23S rRNA links *Chlamydiales*, chloroplasts, and mitochondria. *J. Bacteriol.* **181**:4734–4740.
22. **Ferat, J.-L., and F. Michel.** 1993. Group II self-splicing introns in bacteria. *Nature* **364**:358–361.
23. **Ferat, J.-L., M. Le Gouar, and F. Michel.** 1994. Multiple group II self-splicing introns in mobile DNA from *Escherichia coli*. *C. R. Acad. Sci.* **317**:141–148.
24. **Foley, S., A. Bruttin, and H. Brussow.** 1999. Widespread distribution of a group I intron and its three deletion derivatives in the lysin gene of *Streptococcus thermophilus* bacteriophages. *J. Virol.* **74**:611–618.
25. **Goodrich-Blair, H., V. Scarlato, J. M. Gott, M.-Q. Xu, and D. A. Shub.** 1990. A self-splicing group I intron in the DNA polymerase gene of *Bacillus subtilis* bacteriophage SPO1. *Cell* **63**:417–424.
26. **Goodrich-Blair, H., and D. A. Shub.** 1994. The DNA polymerase genes of several HMU-bacteriophages have similar group I introns with highly divergent open reading frames. *Nucleic Acids Res.* **22**:3715–3721.
27. **Goodrich-Blair, H., and D. A. Shub.** 1996. Beyond homing: competition between intron endonucleases confers a selective advantage on flanking genetic markers. *Cell* **84**:211–221.
28. **Gorbalenya, A. E.** 1994. Self-splicing group I and group II introns encode homologous (putative) DNA endonucleases of a new family. *Protein Sci.* **3**:1117–1120.
29. **Gott, J. M., D. A. Shub, and M. Belfort.** 1986. Multiple self-splicing introns in bacteriophage T4: evidence from autocatalytic GTP labelling of RNA *in vitro*. *Cell* **47**:81–87.
30. **Gott, J. M., A. Zeeh, D. Bell-Pederson, K. Ehrenman, M. Belfort, and D. A. Shub.** 1988. Genes within genes: independent expression of phage T4 intron open reading frames and the genes in which they reside. *Genes Dev.* **2**:1791–1799.
31. **Hensgens, L. A., L. Bonen, M. de Haan, G. van der Horst, and L. A. Grivell.** 1983. Two intron sequences in yeast mitochondrial COX1 gene: homology among URF-containing introns and strain-dependent variation in flanking exons. *Cell* **32**:379–389.

32. Huang, C. C., M. Narita, T. Yamagata, Y. Itoh, and G. Endo. 1999. Structure analysis of a class II transposon encoding the mercury resistance of the gram-positive bacterium *Bacillus megaterium* MB1, a strain isolated from Minimata Bay, Japan. *Gene* **234**:361–369.
33. Huang, Y.-J., M. M. Parker, and M. Belfort. 1999. Role of exonucleolytic degradation in group I intron homing in phage T4. *Genetics* **153**:1501–1512.
34. Jacquier, A., and B. Dujon. 1985. An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell* **41**:383–394.
35. Jurica, M. S., and B. L. Stoddard. 1999. Homing endonucleases: structure, function, and evolution. *Cell. Mol. Life Sci.* **55**:1304–1326.
36. Knoop, V., and A. Brennicke. 1994. Evidence for a group II intron in *Escherichia coli* inserted into a highly conserved reading frame associated with mobile DNA sequences. *Nucleic Acids Res.* **22**:1167–1171.
37. Kowalski, J. C., M. Belfort, M. A. Stapleton, M. Holpert, J. T. Dansereau, S. Pietrokowski, S. M. Baxter, and V. Derbyshire. 1999. Configuration of the catalytic GIY-YIG domain of intron endonuclease I-TevI: coincidence of computational and molecular findings. *Nucleic Acids Res.* **27**:2115–2125.
38. Krasin, F., and F. Hutchinson. 1977. Repair of DNA double-strand breaks in *Escherichia coli*, which requires *recA* function and the presence of a duplicate genome. *J. Mol. Biol.* **116**:81–98.
39. Kuhlmann, U. C., G. R. Moore, R. James, C. Kleanthous, and A. M. Hemmings. 1999. Structural parsimony in endonuclease active sites: should the number of homing endonuclease families be redefined? *FEBS Lett.* **463**:1–2.
40. Kuhse, M. G., R. Strickland, and J. D. Palmer. 1990. An ancient group I intron shared by eubacteria and chloroplasts. *Science* **250**:1570–1573.
41. Kulaeva, O. I., E. V. Koonin, J. C. Wootton, A. S. Levine, and R. Woodgate. 1998. Unusual insertion element polymorphisms in the promoter and terminator regions of the *mucAB*-like genes of R471a and R466b. *Mutat. Res.* **397**:247–262.
42. Lambowitz, A. M., and M. Belfort. 1993. Introns as mobile genetic elements. *Annu. Rev. Biochem.* **62**:587–622.
43. Lambowitz, A. M., M. G. Caprara, S. Zimmerly, and P. S. Perlman. 1999. Group I and II ribozymes as RNPs: clues to the past and guides to the future, p. 451–485. *In* R. F. Gesteland, T. R. Cech, and J. F. Atkins (ed.), *The RNA world*, 2nd ed. (monograph 37). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
44. Landthaler, M., and D. A. Shub. 1999. Unexpected abundance of self-splicing group I introns in the genome of bacteriophage Twort: introns in multiple genes, a single gene with three introns, and exon skipping by group I ribozymes. *Proc. Natl. Acad. Sci. USA* **96**:7005–7010.
45. Lazarevic, V., B. Soldo, A. Dusterhoff, H. Hilbert, C. Mauel, and D. Karamata. 1998. Introns and intein coding sequence in the ribonucleotide reductase genes of *Bacillus subtilis* temperate bacteriophage SP8. *Proc. Natl. Acad. Sci. USA* **95**:1692–1697.
46. Lehman, I. R., and G. A. Pratt. 1960. On the structure of the glycosylated hydroxymethylcytosine nucleotides of T2, T4 and T6. *J. Biol. Chem.* **235**:3254–3259.
47. Loizos, N., G. H. Silva, and M. Belfort. 1996. Intron-encoded endonuclease I-TevII binds across the minor groove and induces two distinct conformational changes in its DNA substrate. *J. Mol. Biol.* **255**:412–424.
48. Martínez-Abarca, F., S. Zekri, and N. Toro. 1998. Characterization and splicing *in vivo* of a *Sinorhizobium meliloti* group II intron associated with particular insertion sequences of the IS630-Tc1/IS3 retroposon superfamily. *Mol. Microbiol.* **28**:1295–1306.
49. Matsuura, M., R. Saldanha, H. Ma, H. Wank, J. Yang, G. Mohr, S. Cavanaugh, G. M. Dunny, M. Belfort, and A. M. Lambowitz. 1997. 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.
50. Michel, F., and D. J. Cummings. 1985. Analysis of class I introns in a mitochondrial plasmid associated with senescence of *Podospora anserina* reveals extraordinary resemblance to the *Tetrahymena* ribosomal intron. *Curr. Genet.* **10**:69–79.
51. Michel, F., and B. Dujon. 1986. Genetic exchange between bacteriophage T4 and filamentous fungi? *Cell* **46**:323.
52. Michel, F., K. Umeson, and H. Ozeki. 1989. Comparative and functional anatomy of group II catalytic introns—a review. *Gene* **82**:5–30.
53. Michel, F., and J. L. Ferat. 1995. Structure and activities of group II introns. *Annu. Rev. Biochem.* **64**:435–461.
54. Mikkonen, M., and T. Alatossava. 1995. A group I intron in the terminase gene of *Lactobacillus delbrueckii* subsp. *lactis* phage LL-H. *Microbiology* **141**:2183–2190.
55. Mills, D. A., D. A. Manias, L. L. McKay, and G. M. Dunny. 1997. Homing of a group II intron from *Lactococcus lactis* subsp. *lactis* ML3. *J. Bacteriol.* **179**:6107–6111.
56. Mills, D. A., L. L. McKay, and G. M. Dunny. 1996. Splicing of a group II intron involved in the conjugative transfer of pRS01 in lactococci. *J. Bacteriol.* **178**:3531–3538.
57. Mueller, J. E., J. Clyman, Y. J. Huang, M. M. Parker, and M. Belfort. 1996. Intron mobility in phage T4 occurs in the context of recombination-dependent DNA replication by way of multiple pathways. *Genes Dev.* **10**:351–364.
58. Mueller, M. W., M. Allmaier, R. Eskes, and R. J. Schweyen. 1993. Transposition of group II intron all in yeast and invasion of mitochondrial genes at new locations. *Nature* **366**:174–176.
59. Mullany, P., M. Pallen, M. Wilks, J. R. Stephen, and S. Tabaqchali. 1996. A group II intron in a conjugative transposon from the gram-positive bacterium, *Clostridium difficile*. *Gene* **174**:145–150.
60. Muscarella, D. E., E. L. Ellison, B. M. Ruoff, and V. M. Vogt. 1990. Characterization of I-PpoI, an intron-encoded endonuclease that mediates homing of a group I intron in the ribosomal DNA of *Physarum polycephalum*. *Mol. Cell. Biol.* **10**:3386–3396.
61. Nauta, A. 1997. Molecular characterization and exploitation of the temperate *Lactococcus lactis* bacteriophage rIt. Ph.D. thesis. University of Groningen, Groningen, The Netherlands.
62. Nikolcheva, T., and S. A. Woodson. 1997. Association of a group I intron with its splice junction in 50S ribosomes: implications for intron toxicity. *RNA* **3**:1016–1027.
63. Ohman-Heden, M., A. Ahgren-Stalhandske, S. Hahne, and B. M. Sjöberg. 1993. Translation across the 5'-splice site interferes with autocatalytic splicing. *Mol. Microbiol.* **7**:975–982.
64. Paquin, B., S. D. Kathe, S. A. Nierzwicki-Bauer, and D. A. Shub. 1997. Origin and evolution of group I introns in cyanobacterial tRNA genes. *J. Bacteriol.* **179**:6798–6806.
65. Paquin, B., A. Heinfing, and D. A. Shub. 1999. Sporadic distribution of tRNA^{ACC} introns among α -purple bacteria: evidence for horizontal transmission and transposition of a group I intron. *J. Bacteriol.* **181**:1049–1053.
66. Parker, M. M., M. Belisle, and M. Belfort. 1999. Intron homing with limited exon homology: illegitimate double-strand-break repair in intron acquisition by phage T4. *Genetics* **153**:1513–1523.
67. Quirk, S. M., D. Bell-Pedersen, J. Tomaszewski, W. Ruger, and M. Belfort. 1988. The inconsistent distribution of introns in the T-even phages indicates recent genetic exchanges. *Nucleic Acids Res.* **17**:301–315.
68. Quirk, S. M., D. Bell-Pedersen, and M. Belfort. 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.
69. Reinhold-Hurek, B., and D. A. Shub. 1992. Self-splicing introns in tRNA genes of widely divergent bacteria. *Nature* **357**:173–176.
70. Roger, A. J., and W. F. Doolittle. 1993. Why introns-in-pieces? *Nature* **364**:289–290.
71. Roman, J., and S. A. Woodson. 1998. Integration of the *Tetrahymena* group I intron into bacterial rRNA by reverse splicing *in vivo*. *Proc. Natl. Acad. Sci. USA* **95**:2134–2139.
72. Roman, J., M. H. Rubin, and S. A. Woodson. 1999. Sequence specificity of *in vivo* reverse splicing of the *Tetrahymena* group I intron. *RNA* **5**:1–13.
73. Romine, M. F., L. C. Stillwell, K.-K. Wong, S. J. Thurston, E. C. Sisk, C. Sensen, T. Gaasterland, J. K. Fredrickson, and J. D. Saffer. 1999. Complete sequence of a 184-kilobase catabolic plasmid from *Sphingomonas aromatica* F199. *J. Bacteriol.* **181**:1585–1602.
74. Rudi, K., and K. S. Jakobsen. 1997. Cyanobacterial tRNA^{Leu}_{UAA} group I introns have a polyphyletic origin. *FEMS Microbiol. Lett.* **156**:293–298.
75. Saldanha, R., G. Mohr, M. Belfort, and A. M. Lambowitz. 1993. Group I and II introns. *FASEB J.* **7**:15–24.
76. Saldanha, R., B. Chen, H. Wank, M. Matsuura, J. Edwards, and A. M. Lambowitz. 1999. RNA and protein catalysis in group II intron splicing and mobility reactions using purified components. *Biochemistry* **38**:9069–9083.
77. Sellem, C. H., G. Lecellier, and L. Belcour. 1993. Transposition of a group II intron. *Nature* **366**:176–178.
78. Semrad, K., and R. Schroeder. 1998. A ribosomal function is necessary for efficient splicing of the T4 phage *thymidylate synthase* intron *in vivo*. *Genes Dev.* **12**:1327–1337.
79. Sharp, P. A. 1991. Five easy pieces. *Science* **254**:663.
80. Shearman, C., J.-J. Godon, and M. Gasson. 1996. Splicing of a group II intron in a functional transfer gene of *Lactococcus lactis*. *Mol. Microbiol.* **21**:45–53.
81. Shub, D. A., M.-Q. Xu, J. M. Gott, A. Zeeh, and L. D. Wilson. 1987. A family of autocatalytic group I introns in bacteriophage T4. *Cold Spring Harb. Symp. Quant. Biol.* **52**:193–200.
82. Shub, D. A., J. M. Gott, M.-Q. Xu, F. B. Lang, F. Michel, J. Tomaszewski, J. Pedersen-Lane, and M. Belfort. 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.
83. Shub, D. A., H. Goodrich-Blair, and S. R. Eddy. 1994. Amino acid sequence motif of group I intron endonucleases is conserved in open reading frames of group II introns. *Trends Biochem. Sci.* **19**:402–404.
84. Weiner, A. M. 1993. mRNA splicing and autocatalytic introns: distant cousins or the products of chemical determinism? *Cell* **72**:161–164.
85. Woodson, S. A., and T. R. Cech. 1989. Reverse self-splicing of the *Tetrahymena* group I intron: implication for the directionality of splicing and for intron transposition. *Cell* **57**:335–345.
86. Woodson, S. A. 1998. Ironing out the kinks: splicing and translation in bacteria. *Genes Dev.* **12**:1243–1247.
87. Xu, M. Q., S. D. Kathe, H. Goodrich-Blair, S. A. Nierzwicki-Bauer, and D. A.

- Shub.** 1990. Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in cyanobacteria. *Science* **250**:1566–1570.
88. **Yang, J., S. Zimmerly, P. S. Perlman, and A. M. Lambowitz.** 1996. Efficient integration of an intron RNA into double-stranded DNA by reverse splicing. *Nature* **381**:332–335.
89. **Yeo, C. C., J. M. Tham, M. W. Yap, and C. L. Poh.** 1997. Group II intron from *Pseudomonas alcaligenes* NCIB 9867 (P25X): entrapment in plasmid RP4 and sequence analysis. *Microbiology* **143**:2833–2840.
90. **Young, P., M. Ohman, M. Q. Xu, D. A. Shub, and B.-M. Sjöberg.** 1994. Intron-containing T4 bacteriophage gene *sunY* encodes an anaerobic ribonucleotide reductase. *J. Biol. Chem.* **269**:20229–20232.
91. **Zhou, L., D. A. Manias, and G. L. Dunny.** 2000. Regulation of intron function: efficient splicing *in vivo* of a bacterial group II intron requires a functional promoter within the intron. *Mol. Microbiol.* **37**:639–652.
92. **Zimmerly, S., H. Guo, P. S. Perlman, and A. M. Lambowitz.** 1995. Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* **82**:1243–1247.
93. **Zimmerly, S., H. Guo, R. Eskes, J. Yang, P. S. Perlman, and A. M. Lambowitz.** 1995. A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell* **83**:529–583.