Cytogenet Genome Res 110:491–499 (2005)
DOI: 10.1159/000084982

Cytogenetic and Genome Research

Retrons, msDNA, and the bacterial genome

B.C. Lampson,[a] M. Inouye,[b] S. Inouye[b]

[a] Department of Health Sciences, East Tennessee State University, Johnson City, TN (USA); [b] Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, NJ (USA)

Manuscript received 9 September 2003; accepted in revised form for publication by J.-N. Volff 16 October 2003.

Abstract. Retrons are distinct DNA sequences that code for a reverse transcriptase (RT) similar to the RTs produced by retroviruses and other types of retroelements. Retron DNAs are commonly associated with prophage DNA and are found in the genomes of a wide variety of different bacteria. The retron RT is used to synthesize a strange satellite DNA known as msDNA. msDNA is actually a complex of DNA, RNA, and probably protein. It is composed of a small, single-stranded DNA, linked to a small, single-stranded RNA molecule. The 5′ end of the DNA molecule is joined to an internal guanosine residue of the RNA molecule by a unique 2′-5′ phosphodiester bond. msDNA is produced in many hundreds of copies per cell, but its function remains unknown. Although retrons are absent from the genome of most members of a population of related bacteria, retrons may not be entirely benign DNAs. Evidence is beginning to suggest that retron elements may produce small but potentially significant effects on the host cell. This includes the generation of repeated copies of the msDNA sequence in the genome, and increasing the frequency of spontaneous mutations. Because these events involve the retron RT, this may represent a source of reverse transcription in the bacterial cell. Thus, the process of reverse transcription, a force that has profoundly affected the content and structure of most eukaryotic genomes, may likewise be responsible for changes in some prokaryotic genomes.

An incredible amount of DNA, as much as 40%, of the human genome appears to be derived from the reverse transcription of RNA with the subsequent incorporation of these sequences into the genome by retrotransposition events (Kazazian, 2000; Eickbush, 2002). A reverse transcriptase (RT) responsible for perhaps most of these retrotransposition events is encoded by a transposable element called L1, a non-LTR type retrotransposon. Over time hundreds of thousands of copies, mostly truncated copies, of the L1 element have been produced and incorporated into the human genome. In addition, the RT from L1 elements may also be responsible for the accumulation of hundreds of thousands of short repeated sequences such as Alu sequences and SINE elements (Kazazian, 2000; Weiner, 2002). The presence of an active L1-RT may have also produced other changes in the human genome such as the formation of processed pseudogenes, disease causing insertions, and recombinational shuffling of DNA to produce new genes (Kazazian, 2000). These profound changes in the content and structure of the human genome, presumably caused by retrotransposition, have been similarly observed in many other eukaryotic organisms from mammals to insects to plants.

Has retrotransposition, likewise, changed the genomes of earth's oldest organisms, the prokaryotes? For many years after its discovery in animal viruses, RT was thought to be absent from the prokaryotes (Temin, 1989). In fact, RT-encoding elements have now been found in a wide variety of different bacteria and fall into two basic types; retrons and group II introns. The first discovery of an RT-encoding gene in bacteria came from the study of retron elements (Lampson et al., 1989b; Lim and Maas, 1989). Retrons are defined by their unique ability to produce an unusual satellite DNA known as msDNA. Group II introns are DNAs that interrupt genes found in organellar genomes (mitochondria and chloroplasts) of some eukaryotes and encode self-splicing intron RNA (Belfort et al., 2002). They have also been found in many bacterial species as well (Martinez-Abarca and Toro, 2000). With the presence of these two RT elements, has retrotransposition occurred in bacterial genomes? And are the forces of reverse transcription, that have so profoundly affected the genomes of eukaryotes, also at work on the genomes of prokaryotes? These are interesting questions given the widespread

B.C.L. was supported by a grant from Takara Bio Inc. and from the office of research at East Tennessee State University. S.I. and M.I. are supported by a grant (GM 44012) from the NIH.

Request reprints from: Dr. Sumiko Inouye, Department of Biochemistry
Robert Wood Johnson Medical School
University of Medicine and Dentistry of New Jersey
Piscataway, NJ 08854 (USA); telephone: +1-732-235-4161
fax: +1-732-235-4559; e-mail:sinouye@waksman.rutgers.edu

1424–8581/05/1104–491$22.00/0

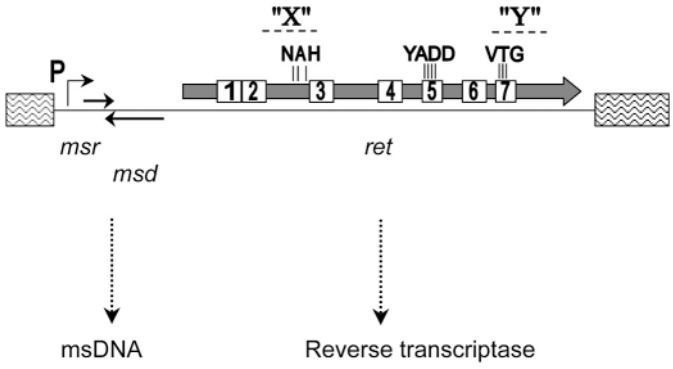

Fig. 1. Functional elements encoded in retron DNA. A retron element is composed of about 2 kbp of unique DNA (thin line) usually inserted into the chromosome (flanking boxes filled with wavy lines). Retron DNA codes for a promoter sequence (bent arrow labeled with a P) that controls the transcription of a simple operon composed of three loci: *msr*, *msd*, and *ret*. The "gene" *msr* codes for the RNA molecule found in msDNA and *msd* codes for the DNA chain found in msDNA. Adjacent to *msr-msd* is a long ORF encoding a RT consisting of usually about 300–400 amino acid residues. The gene encoding the RT is designated *ret* (large arrow). Like all RTs the retron protein contains seven regions of conserved amino acids (small open squares labeled 1–7) including the conserved YADD sequence associated with the catalytic core of the polymerase enzyme (letters shown above region 5) (Steitz, 1999). There are two regions within the retron protein that contain conserved amino acids found only in retron RTs (designated "X" and "Y"). They contain conserved amino acids N, A, and H in region "X" and the sequence VTG in region "Y"(Inouye et al., 1999). Some retron elements may encode a second ORF of unknown function (see also Fig. 4).

occurrence and possible long-term association of bacterial cells with active RT-producing elements. Recent evidence suggests that retron elements may produce some effects on the host organism, possibly via retrotransposition events. An overview of retron elements will be presented here along with a discussion of the possible effects retrons may produce on their host genomes.

Retron DNA

Because it is not known whether retrons are mobile DNAs, these elements have, instead, been defined by their ability to synthesize an unusual satellite DNA known as msDNA (Dhundale et al., 1987). Despite considerable investigation, the function of this strange satellite DNA remains unknown. Based on this activity, retrons are distinct DNA sequences of about 2 kilobasepairs (kbp) in size, and are found inserted into the bacterial chromosome or as part of a large prophage DNA element.

All retron DNAs contain at least one open reading frame (ORF) that codes for a protein similar to eukaryotic (viral) RTs and is designated *ret* (Fig. 1). A careful comparison of the amino acid sequence of the retron ORF with eukaryotic RTs clearly reveals seven regions of conserved amino acid residues shared by all RTs (Xiong and Eickbush, 1990). These conserved amino acids correspond to conserved structures within the RT protein; the so called palm, fingers, and thumb regions found in the crystal structure of HIV-RT (Steitz, 1999). Besides these seven regions, retron RTs appear to have two additional blocks of unique amino acid residues not found in other RTs and are designated regions "X" and "Y" (Fig. 1). One explanation for these unique regions of conserved amino acids is that amino acid residues at the C-terminus of the RT protein, that includes region "Y", appear to be unique to individual RTs and crucial for specific recognition and binding to the template-primer RNA used for the synthesis of msDNA (Inouye et al., 1999). The unusual priming mechanism used to synthesize msDNA thus appears to be a unique function of retron RTs (see msDNA below).

Aside from RT activity, no other functional domains are known to exist within the retron RT sequence. For example, RNase H activity may play a role in the synthesis of msDNA, but no RNase H activity is known to be associated with retron RTs and is instead provided by enzymes produced by the host cell (Lima and Lim, 1995; Shimamoto et al., 1995a). Other functional domains such as an endonuclease or maturase region are also absent from retron RT-ORFs.

In addition to the RT-ORF, some retron elements also code for a second ORF. For example, the retron element designated Ec73 from *Escherichia coli* contains a 316 amino acid ORF located just upstream of the RT-ORF (Sun et al., 1991). The function of this second ORF is unknown. Aside from the RT-ORF, all retron elements contain two additional loci ("genes") designated *msd* and *msr*. These two loci are required to synthesize msDNA and are situated in opposite directions such that their 3′ ends overlap by several bases (Fig. 1). The sequence of nucleotides encoded in *msd* correspond to the DNA molecule found in msDNA and that of *msr* corresponds to the molecule of RNA linked to msDNA (see msDNA below). Because msDNA is not an autonomously replicating satellite DNA, the genes *msr*, *msd*, and *ret* (RT) essentially form an operon required to synthesize msDNA. Indeed, experimental evidence indicates that the retron genes *msr*, *msd*, and *ret* are transcribed together as one long mRNA and, as described below, msDNA is then produced by a reverse transcription mechanism. Transcription of this retron operon appears to be controlled by a promoter found within the retron DNA just upstream of *msr* (Herzer et al., 1992).

The unusual structure of msDNA

msDNA was first discovered in a myxobacterium, *Myxococcus xanthus* and is an abbreviation for multi-copy single-stranded DNA. Indeed, the msDNA molecule produced by this bacterium, designated msDNA-Mx162, contains a single-stranded DNA molecule of 162 nucleotides and as many as 500 to 700 copies are produced per cell (Yee et al., 1984). However, from simple experiments, such as the instability of 5′end-labeled molecules to exposure to sodium hydroxide, and the electrophoretic mobility change of the molecule after treatment with RNase, it was quickly realized that RNA is associated with msDNA (Furuichi et al., 1987). In fact, RNA is covalently linked to the single-stranded DNA molecule by a unique 2′-5′ phosphodiester linkage (Dhundale et al., 1987).

An msDNA, designated msDNA-Ec73, from *E. coli*, illustrates the unusual branch structure of msDNA as well as other conserved structures of this satellite molecule (Fig. 2). First, msDNA-Ec73 contains a single-stranded DNA of 73 nucleotides that folds into a long, stable, stem-loop structure. Of the dozen or so msDNAs that have been characterized from bacteria all contain a single-stranded DNA (from 48 to 165 nucleotides) that is postulated to fold into a stable secondary structure. Covalently

bonded to the 5′ end of this DNA molecule is a single-strand of RNA (boxed sequence in Fig. 2). That is, at a specific guanosine residue in the middle of the RNA strand, the DNA molecule is joined to the 2′ position of this guanosine residue of the RNA. Thus, msDNA is a unique molecule containing a 2′-5′ phosphodiester linkage between the RNA and DNA strands. The branch linkage in the RNA molecule always occurs at a specific internal G residue usually within an AGC sequence. The last conserved structure in msDNA is a short region of base pairing between the 3′ end of the DNA strand and the 3′ end of the RNA chain (Fig. 2). Beyond these conserved secondary structures, the primary nucleotide sequence of both the DNA and RNA molecules is generally highly variable between any two different msDNAs. There are groups of related retrons that produce msDNAs with some degree of primary sequence similarity (see below).

The conserved secondary structures in msDNA, the branch linkage, the folded stem-loops, and the RNA-DNA hybrid region are present because of the way msDNA is synthesized (see below). However, after synthesis is complete, in some retrons msDNA is further modified. For example, msDNA-Mx162 from *M. xanthus* is initially synthesized with a 77-base RNA molecule joined to the 162-base DNA molecule. However, this precursor form is then converted to a smaller mature form of msDNA by the removal of 16 nucleotides from the 5′ end of the RNA chain (Furuichi et al., 1987b). For some retron elements post-production processing of msDNA results in the complete removal of the entire RNA molecule including the branch structure. For example, in the case of msDNA-Ec83 (from *E. coli*) an endonucleolytic cleavage of the DNA molecule between the fourth and fifth nucleotides removes the 84-base RNA chain plus four deoxyribonucleotides joined to the branched G residue (Lim, 1992; Kim et al., 1997). This results in a mature msDNA composed only of a single-stranded DNA molecule. There is some speculation that this msDNA may have autocatalytic activity or the retron RT may be responsible for the cleavage of the DNA molecule during the processing event. However, the source of this endonuclease activity remains unknown. There appears to be a group of phylogenetically related retron elements that produce these mature, RNA-free msDNAs. They include the retrons Ec78, and Ec83 from *E. coli* (Lima and Lim, 1997), Vc 95 from *Vibrio cholerae* (Shimamoto et al., 1999; Ahmed and Shimamoto, 2003), and St85 from *Salmonella typhimurium*; (Ahmed and Shimamoto, 2003; Matiasovicova et al., 2003). The msDNAs produced by these retrons also appear to share some degree of primary nucleotide sequence similarity in their DNA and RNA molecules.

One final aspect of the structure of msDNA should not be overlooked. msDNA was originally discovered because it appears as a small but prominent satellite DNA band after gel electrophoresis of total DNA extracted from *M. xanthus* cells (Yee et al., 1984). However, in all likelihood msDNA probably exists apart from the chromosome as a large molecular weight nucleoprotein complex composed of multiple copies of msDNA bound to protein components (Lampson et al., 1990). Certainly, the retron RT is very likely the major protein bound to msDNA in this complex. When RT from retron-Ec67 was purified it co-fractionates with msDNA as a large molecular weight complex (Lampson et al., 1990). Other proteins, produced by the host cell, may also bind to the msDNA complex. For example, RNase H produced by the host cell may bind to msDNA, at least transiently, because it appears to be involved in the synthesis of msDNA (Lima and Lim, 1995; Shimamoto

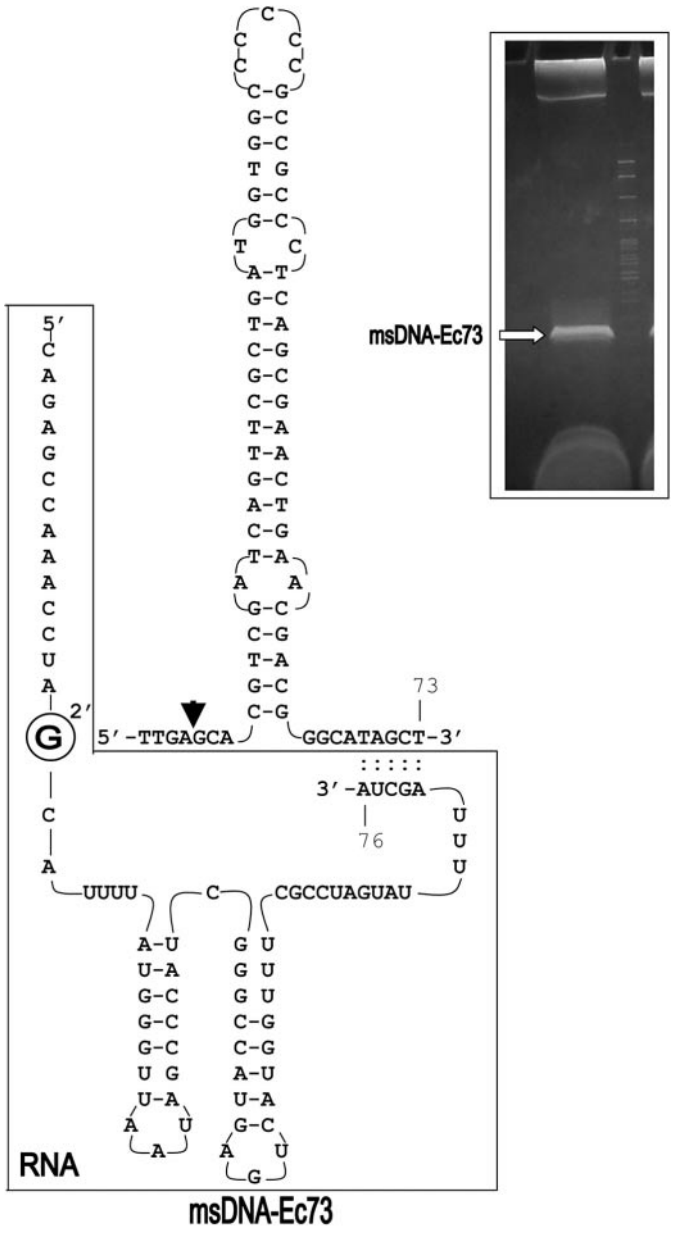

Fig. 2. The structure of msDNA. msDNA, for example msDNA-Ec73, is a small molecular weight, satellite DNA (white arrow in inset photo) that is easily observed after gel electrophoresis of plasmid DNA extracts from *E. coli* (Sun et al., 1991). This msDNA is composed of a 76-nucleotide RNA molecule (boxed sequence) joined to a 73-nucleotide single-stranded DNA molecule. The 5′ end of the DNA chain is joined, via a 2′,5′ phosphodiester bond, to a specific guanosine residue of the RNA molecule (circled G). Both the RNA chain and the DNA chain fold-up into stable stem-loop structures. Some msDNAs are altered (processed) after they are synthesized (see text for details). For example, about 16 nucleotides are clipped off the 5′ end of the RNA chain in msDNA-Mx162 from *M. xanthus* resulting in a mature form of the msDNA with a smaller RNA molecule. Likewise, for some msDNAs the entire RNA chain is removed (as well as 4 nucleotides from the 5′ end of the DNA molecule) by an endonucleolytic cleavage at the 5′ end of the DNA strand (→).

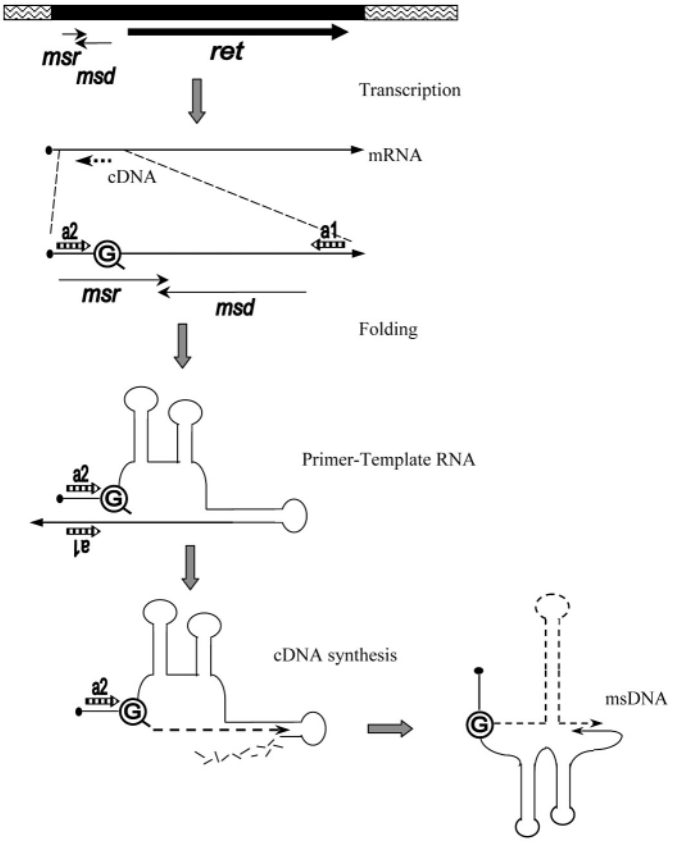

Fig. 3. Synthesis of msDNA. After transcription of the retron operon *msr-msd-ret*, a small part of the mRNA serves as both the template and primer for the synthesis of a cDNA copy of the *msd* region by reverse transcription. Folding of the primer-template RNA into a secondary structure allows the 2′-OH group of a specific branching G residue (G in a circle with the 2′-OH group shown by a short line) to serve as a primer to initiate cDNA synthesis by the retron RT. After synthesis of cDNA (dashed line) is complete, part of the RNA template remains joined to the 5′ end of the cDNA molecule to yield msDNA. See text for more detail.

et al., 1995a) (see below). In addition mis-match repair proteins like MutS, produced by the host cell, may also recognize and bind to msDNA (Maas et al., 1994). This may have important implications for the host cell and will be discussed below.

The synthesis of msDNA

msDNA is essentially a cDNA copy, produced by reverse transcription, of a short region of an mRNA template. For msDNA, however, part of the RNA template is not reverse transcribed and remains associated with the cDNA after completion of reverse transcription. The unique 2′-5′ linkage in msDNA is produced by the retron RT to initiate cDNA synthesis (Shimamoto et al., 1995b). Here, the 2′-OH group of a specific guanosine residue in the RNA template is used to prime the start of cDNA polymerization (Hsu et al., 1989). Thus, a region of the mRNA transcript encoding the *msr-msd* loci serves as both a primer and template for cDNA synthesis by RT. The development of a fully in vitro system to study the production of msDNA has allowed detailed mechanisms to be revealed and these have been described and reviewed elsewhere (Lampson et al., 2001; Yamanaka et al., 2002). A brief overview of the synthesis of msDNA is discussed here.

Before synthesis of msDNA can begin an RNA transcript of the *msr-msd* region of retron DNA must be made to serve as the primer and template for reverse transcription. This primer-template RNA is derived from the longer mRNA transcript of the retron operon *msr-msd-ret* described earlier (Fig. 3). Crucial to the start of cDNA synthesis is the ability of the RNA template to fold into a specific secondary structure. This specific secondary structure allows the primer-template RNA to initiate cDNA synthesis by the retron RT. Folding of the RNA molecule is mediated by intra-molecular base pairing between two inverted repeat sequences in the RNA, designated a1 and a2 (Fig. 3). The stem structure thus formed, along with other stem-loops, apparently juxtapose a specific internal guanosine residue (called the branching G) within the folded RNA so that the 2′-OH group of this branching G can serve as a primer to start cDNA synthesis (Hsu et al., 1989). The retron RT recognizes and binds to a specific stem-loop structure in the primer-template RNA along with the properly positioned branching G at the end of the a1-a2 stem structure (Fig. 3) (Inouye et al., 1999). This allows the RT to utilize the 2′-OH of the branching G (circled G in Fig. 3) as a primer to initiate cDNA synthesis. After formation of the 2′-5′ phosphodiester bond with the first deoxyribonucleotide (dNTP), further template directed incorporation of dNTPs occurs in a conventional manner to continue synthesis of the cDNA. In close concert with the extension of the cDNA molecule is the removal of the lagging RNA template by an RNase H activity, apparently supplied by the host cell (Shimamoto et al., 1995a). At a specific location on the RNA template cDNA polymerization and RNase H mediated removal of the RNA template stop, resulting in the msDNA molecule observed in cell extracts. That is, a short cDNA molecule covalently joined to the remainder of the RNA molecule that served as a primer-template for its synthesis.

Retrons and the bacterial genome

Despite considerable investigation, retrons remain a puzzle. It is not known if they are mobile elements, and so they are defined by the only activity ascribed to them: the production of msDNA. The synthesis of msDNA appears to use a unique and elaborate mechanism, requires a special type of RT, is produced in large numbers, and in some cases undergoes post-production processing. Yet its function remains unknown. A knock-out mutation of the retron-RT gene, *ret*, with an antibiotic resistance marker eradicates production of msDNA but otherwise has no discernible effect on the viability of the host cell (Dhundale et al., 1988; Inouye et al., 1990). Despite its obscure function, retrons and msDNA may not be entirely benign DNAs. Recent evidence suggests that retrons and their product, msDNA, may produce subtle yet potentially important changes to the host genome.

Prevalence of retrons in bacterial genomes

Retron elements appear to be present in the genomes of a wide variety of different eubacteria from at least 7 different major taxa (Table 1). This is based on two different criteria for identifying retrons in bacteria. The first is the detection of bacteria that produce msDNA and therefore must contain a retron.

Table 1. Bacteria that contain a retron

Bacterial species	Bacterial taxon[a]	Retron designation[b]	Produces msDNA[c]	Presence of RT gene[d]	Reference
Escherichia coli	Gamma Proteobacteria	Ec48	yes	yes	Mao et al., 1997
		Ec67	yes	yes	Lampson et al., 1989b
		Ec73	yes	yes	Sun et al., 1991
		Ec78	yes	ND[e]	Lima & Lim, 1997
		Ec83	yes	yes	Lim, 1992
		Ec86	yes	yes	Lim & Maas, 1989
		Ec107	yes	yes	Herzer et al., 1992
Salmonella typhimurium	Gamma-Proteobacteria	St85	yes	yes	Ahmed & Shimamoto, 2003
Salmonella enteritidis	Gamma-Proteobacteria	ND	yes	yes	Rychlik et al., 2001
Vibrio cholerae	Gamma-Proteobacteria	Vc95	yes	yes	Shimamoto et al., 1999
Vibrio parahaemolyticus	Gamma-Proteobacteria	Vc96	yes	yes	Ahmed & Shimamoto, 2003
Klebsiella pneumoniae	Gamma-Proteobacteria	ND	yes	ND	Rice et al., 1993
Proteus mirabilis	Gamma-Proteobacteria	ND	yes	ND	Rice et al., 1993
Xanthomonas campestris	Gamma-Proteobacteria	ND	ND	yes	Rest & Mindell, 2003
Rhizobium sp.	Alpha-Proteobacteria	ND	yes	ND	Rice et al., 1993
Bradyrhizobium sp.	Alpha-Proteobacteria	ND	yes	ND	Rice et al., 1993
Ralstonia metallidurans	Beta-Proteobacteria	ND	ND	yes	Rest & Mindell, 2003
Myxococcus xanthus	Delta-Proteobacteria	Mx65, Mx162	yes	yes	Dhundale et al., 1987
Stigmatella aurantiaca	Delta-Proteobacteria	Sa163	yes	yes	Hsu et al., 1992
Nannocystis exedens	Delta-Proteobacteria	Ne144	yes	yes	Lampson et al., 2002
Geobacter sulfurreducens	Delta-Proteobacteria	ND	ND	yes	This work
Trichodesmium erythraeum	Cyanobacteria	ND	ND	yes	Rest & Mindell, 2003
Nostoc punctiforme	Cyanobacteria	ND	ND	yes	Rest & Mindell, 2003
Nostoc sp.	Cyanobacteria	ND	ND	yes	Rest & Mindell, 2003
Staphylococcus aureus	Gram positive	ND	ND	yes	Matiasovicova et al., 2003
Fusobacterium nucleatum	Bacteroides-Flavobacter	ND	ND	yes	Matiasovicova et al., 2003
Flexibacter elegans	Bacteroides-Flavobacter	ND	yes	ND	Dhundale et al., 1985

[a] Taxonomic groups are based on Woese, 1987.
[b] Retrons are designated with two letters based on the genus and species, respectively, of their host organism followed by a number based on the size of the DNA molecule of the msDNA they produce.
[c] This is based on experimental detection of msDNA from cell extracts.
[d] This is based on DNA sequence determination of the retron or the presence of the DNA sequence in microbial genome DNA databases.
[e] ND, not determined

The second method is based on the presence of a retron-type RT gene in the genome, discovered from the search of a DNA database like GenBank (Matiasovicova et al., 2003, Rest and Mindell, 2003). This second approach has recently revealed a retron-like RT gene in a member of the Archaea as well (Rest and Mindell, 2003).

Although retrons are widely distributed among the bacteria, within a population of bacteria of the same species (or closely related species) retrons tend to be rare. A good illustration of this is the study of retrons from *E. coli*. Seven distinct retron elements, with little or no similarity between them, have been found from various *E. coli* strains (Table 1). An analysis of a reference population of *E. coli* strains known as the ECOR collection (Selander et al., 1987) indicated that only 11 out of 75 (15%) ECOR strains produces msDNA (Herzer et al., 1990). The rare occurrence of retron elements was similarly noted among collections of strains from the rhizobia (16%), *Klebsiella* (5%), *Proteus* (17%), and *Salmonella* (6%) (Rice et al., 1993). In addition, one retron in particular appears to predominate among the ECOR collection; retron Ec107 is found in eight of the 75 ECOR strains (Kawaguchi et al., 1992). Based on Southern hybridization experiments only one copy of retron Ec107 is present in the genome of these bacteria.

Although retrons appear to be quite rare in the ECOR population, mobile genetic elements like group II introns and insertion sequences are quite common. For example, 5 different types of group II introns are found among the ECOR collection with all 75 ECOR strains (100%) containing at least one copy of intron E.c.I4 and some strains like ECOR 38 contain as many as 15 copies of this intron element (Dai and Zimmerly, 2002). Insertion sequences (IS) occur even more frequently among the ECOR collection. For example, six different IS elements are found among the ECOR population with 96% of the strains containing at least one IS element (Sawyer et al., 1987). Most of the ECOR strains contain several copies of IS elements such as strain ECOR 36 which contains 25 copies of IS1, 6 copies of IS2, 14 copies of IS4, 3 copies of IS5, and 2 copies of IS30.

Two factors could explain the rare occurrence of retron elements among *E. coli* strains. First, retron DNA may be a very recent addition to the *E. coli* genome and secondly retrons transpose to a new location very rarely, if at all. Indeed, evidence indicates that retrons may have recently been incorporated into

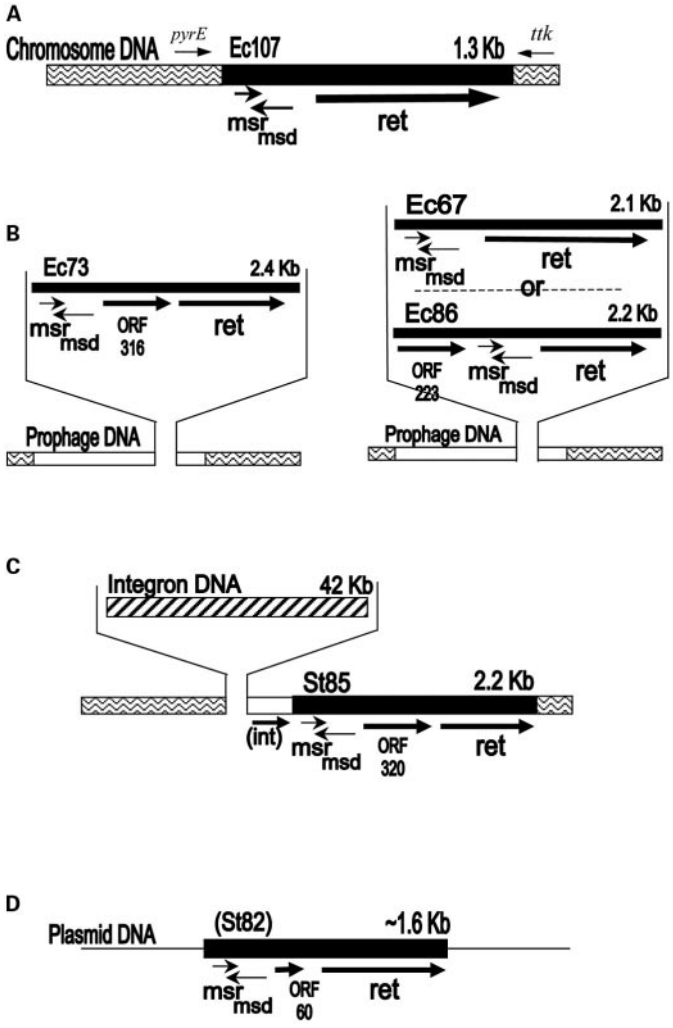

Fig. 4. Four examples of retron insertion sites. (**A**) Retron-Ec107 DNA (black bar) for example, is inserted into an intergenic region of the *E. coli* chromosome (bars filled with wavy lines) between the *pyrE* and *ttk* genes (Herzer et al., 1992). (**B**) Most retron elements, at least from *E. coli*, appear to have inserted into prophage DNA (open bar). For example, retron-Ec73 is located in a P4-like prophage DNA integrated into the *selC* gene on the *E. coli* chromosome (Sun et al., 1991). Two different retrons, Ec67 and Ec86 appear to have inserted, on separate occasions, into the same site of a phage 186-related prophage element in the *E. coli* chromosome (Hsu et al., 1990; Lim, 1991, respectively). (**C**) Retron-St85 appears to be inserted by a complex event into the *Salmonella typhimurium* chromosome (Boyd et al., 2000; Ahmed and Shimamoto, 2003). Adjacent to the retron DNA (black bar) is a phage-like integrase pseudogene (open bar with arrow labeled int) suggesting that the retron was perhaps part of a larger prophage element. Integrated next to this DNA is a large 42-kbp "genome island" (hatched bar) containing an integron element (see also text). (**D**) A retron, (its designation was not determined by Rychlik et al., but is tentatively assigned here as St82) from *Salmonella enteritidis*, appears to be integrated into plasmid DNA (Rychlik et al., 2001).

the genome of a few strains of *E. coli*. This is based on a pattern of atypical codons (for *E. coli* genes) found in the retron-encoded RT gene of Ec107 and the lower than expected nucleotide sequence variability among individual Ec107 elements in the ECOR collection (Kawaguchi et al., 1992; Herzer et al., 1992).

Since almost all RT genes appear to function in either the replication or (and) transposition of various genetic elements it seems likely that retron elements may also be mobile DNAs; but this remains unproven for any retron. So how did some *E.coli* bacteria come to acquire these msDNA producing elements? It appears that most of the retrons found in *E. coli* are inserted into prophage DNA (see retron insertion sites below) found in the chromosome. One prominent exception to this rule is the *E. coli* retron Ec107. Bacteriophages may have been a likely vehicle for the delivery of retron DNA into the *E. coli* chromosome. Indeed, a retron-containing prophage designated φR73, which is related to phage P4, can be induced to excise from the chromosome and form infectious virions (Inouye et al., 1991). The phage particles can then infect a new *E. coli* host strain and deliver the retron to the chromosome via a prophage and render the lysogenized host capable of producing msDNA.

Although retrons appear to be rare additions to the genomes of most bacterial groups, one exception is found in the myxobacteria. In a population of 20 different strains of *M. xanthus* all appeared to produce an msDNA the same as Mx162 (Dhundale et al., 1985; Lampson et al., 1991). All 10 genera of the myxobacteria produce msDNA and the retron-RT genes that have been characterized from different genera of the mxyobacteria appear to form a phylogenetically related group (Rice and Lampson, 1995; Lampson et al., 2002). These features along with other evidence like codon usage and nucleotide sequence divergence suggest that retron DNA has resided in the genome of the myxobacteria for perhaps as long as 150 million years when the various species of today all shared a common ancestor (Rice and Lampson, 1995).

Retron insertions into the genome

For a small number of mostly *E. coli* retrons, the precise junction of retron DNA with the host chromosome has been determined and can be used to analyze the insertion site. Characterization of the insertion site can provide important clues as to how retrons were acquired by the bacterial genome and the nature of the insertion event.

As mentioned above, most of the *E. coli* retrons so far examined appear to have inserted into prophage DNA found on occasion in the chromosome (Fig. 4). A good example is retron Ec73 found in a clinical strain of *E. coli* (Sun et al., 1991). This 2.4-kbp retron DNA is inserted into the right end of a 12.7-kbp prophage element that lies at 82.3 min on the chromosome of *E. coli* strain Cl-23 (Fig. 4). This prophage DNA resembles the phage P4, which is a parasitic phage of the P2 family (Halling and Calendar, 1990). In another example, apparently as a result of two independent events, two different retron DNAs, Ec67 and Ec86, have inserted into precisely the same site of a 34-kbp prophage DNA. This prophage sequence, which resides at 19 min on the *E. coli* chromosome, closely resembles phage 186, also a member of the P2 family of phages (Hsu et al., 1990; Lim, 1991). In each case, retron Ec67 and Ec86 appear to have inserted into a region between the *cos* site and ORF 341 (a DNA packaging gene) of the phage DNA. About 180 bp of phage DNA was apparently lost upon insertion of the retron element (Lim, 1991; Dodd and Egan, 1996).

An interesting exception to the *E. coli* retrons that are associated with prophage DNA is the retron Ec107 (Herzer et al., 1992). This retron is found inserted into an intergenic region of the *E. coli* chromosome between the *pyrE* and *ttk* genes. Appar-

ently, a 34-bp sequence of intergenic DNA was removed when the retron inserted into this site. The mechanism used by the retron to insert into this site remains unclear. There does not appear to be any duplication of a target site sequence flanking the retron DNA, as is found in the case of IS elements. In addition, there is no significant DNA sequence homology between this intergenic region in *E. coli* K-12, which lacks the retron DNA, and DNA sequences at the termini of the retron. This suggests that homologous recombination was not involved in the insertion event.

Some information about the insertion site is known for only two other retron elements outside of *E. coli*. These are two recently described retrons from *Salmonella*. Retron St85 appears to produce an RNA-free type msDNA and is found adjacent to the *yidY* gene in the chromosome of *S. typhimurium* (Ahmed and Shimamoto, 2003; Matiasovicova et al., 2003). The percentage of G plus C nucleotides in the retron DNA is significantly lower than the surrounding *Salmonella* chromosome sequences, suggesting a recent acquisition into the chromosome. In addition, immediately upstream of the retron DNA is a pseudogene encoding a protein that resembles a phage-type integrase (Boyd et al., 2000; Ahmed and Shimamoto, 2003). The pseudogene plus the retron DNA are flanked by direct sequence repeats suggesting that this may have been, at one time, a larger prophage element (Fig. 4). More interesting is the corresponding region of the chromosome found in multi-antibiotic resistant strains of *S. typhimurium* (Boyd et al., 2000). Here, positioned upstream of the retron (plus the pseudogene DNA) is a large 42-kbp "genome island" DNA (Fig. 4). This large genome island (SgiI) appears to contain an integron element carrying several antibiotic resistance genes. It is intriguing that a retron, producing an active RT, is found in close proximity to an integron on the chromosome. There has been some speculation that the gene cassettes carrying antibiotic resistance found in integrons are produced by reverse transcription of mRNA (Recchia and Hall, 1997). It remains to be shown, of course, whether retrons play any role in integron gene cassettes. The other *Salmonella* retron was discovered in a strain of *S. enteritidis* and is carried on a small, 4-kbp plasmid that may also be associated with phage DNA since this plasmid confers phage resistance to the host cell (Rychlik et al., 2001).

msDNA and host mutations

Most retrons produce an msDNA molecule that, due to extensive intramolecular base pairing within the DNA strand, can fold into a long, double-stranded, stem-loop structure (Fig. 2). For some msDNAs, however, like Ec86 and Ec73, the double-stranded stem may occasionally have a mis-matched base pair (Fig. 2). Experiments have shown that if retrons like Ec86 or Ec73 are over-expressed on a multicopy plasmid, such that thousands of copies of msDNA are produced, these msDNAs can be mutagenic to the host cell (Maas et al., 1994, 1996; Mao et al., 1996). For example, overexpression of retron Ec86 (or Ec73) on a plasmid in a *lac*-indicator strain caused an increase in the number of frameshift mutations and substantially increased the frequency of lac+ revertants. However, the increased frequency of these mutations caused by the overexpression of msDNA can be suppressed if the same host cell contains a plasmid that overexpresses the repair protein MutS. It therefore appears that the mutagenic effect of these msDNAs is due to binding of repair proteins, especially MutS produced by the host cell, to mismatched base pairs in the msDNA molecule. Because so many copies (several thousand) of msDNA are produced when over-expressed, a large proportion of the host MutS repair protein is bound-up in msDNA and thus sequestered away from its normal functions. Indeed, when the mismatched base pairs in msDNA-Ec73 are removed by genetic engineering, this modified msDNA is no longer mutagenic (Mao et al, 1996). msDNAs like Ec78 that naturally have no mis-matched base pairs in their stem structure are, likewise, presumed to be non-mutagenic.

An additional effect is seen on the host cell when msDNA with mis-matched base pairs is over-expressed. Here an increase in the frequency of recombination between donor and recipient DNAs during inter-species matings of *Salmonella* and *E. coli* is observed (Maas et al., 1996). Again, this is due to titrating out the MutS repair protein by the msDNA molecule. The MutS protein will remove donor DNA that differs significantly from the recipient (host) DNA due to base pair mismatches and therefore reduce the level of cross-species recombination. Again, it is presumed that this effect does not occur with msDNAs that naturally do not contain mismatched base pairs.

It should be noted that these effects on the frequency of host mutations and recombination are only observed when msDNA is overexpressed. These effects are not seen in the natural situation where a host cell contains a single copy of the retron on the chromosome and is expressed from the retron's native promoter. However, there may be situations where the presence of a retron that produces mutagenic msDNA could influence the host cell. For example, studies have shown that a significant depletion in the levels of MutS, and thus a greatly reduced capacity to carry out mis-match DNA repair, occur in *E. coli* cells when they enter stationary phase (Feng et al., 1996). In addition, experiments with retron Ec107 indicate a marked increase in expression from the retron's promoter when cells are in stationary phase (Herzer, 1996). Thus, the presence of a retron element may affect the rate of mutation and DNA recombination and possibly influence the host cell's ability to evolve under stress conditions.

Retrons and repetitive DNA

As far as is known retrons appear to exist as a single-copy DNA element in the bacterial chromosome or as part of a prophage. There are some examples of bacterial strains that carry two different retron elements each, however, at a single copy. Thus, when msDNA is used as a hybridization probe it will hybridize to a single restriction (endonuclease) fragment of the bacterial chromosome corresponding to the single copy of the retron encoded *msd* gene. It was quite interesting when several different restriction fragments from the chromosome of *Nannocystis exedens* (a myxobacterium) hybridized to msDNA-Ne165 that was radiolabeled for use as a hybridization probe (Lampson and Rice, 1997). Further analysis revealed that numerous partial copies of the msDNA *(msd)* sequence appeared to be dispersed throughout the host chromosome. The repeat sequences varied in size from 22 to 56 nucleotides and represented only part of the total 165 nucleotides present in the msDNA-Ne165 molecule. In

another example of a partial copy of msDNA, again from the chromosome of *N. exedens*, this repeat sequence contained the last 82 nucleotides of the DNA strand of msDNA-Ne144 (Lampson et al., 2002). In addition, this partial copy of msDNA was found to overlap the end of a long ORF of unknown function. The last ten codons plus the termination codon of this ORF, therefore, reside within the msDNA sequence. It is unusual for a repetitive sequence to lie within an ORF (Ogata et al., 2000) since most of the known families of repeated sequences in bacteria, such as REP and ERIC, are found in extragenic DNA in the bacterial chromosome (Lupski and Weinstock, 1992).

How did several partial copies of the msDNA sequence become incorporated into the bacterial chromosome? Although the mechanism is unknown, and conventional mechanisms involving recombination can not be discounted, it appears more likely that these partial copies of msDNA have been generated by reverse transcription. A mechanistic model has been proposed (Lampson and Rice, 1997) that could explain how these repeat sequences are formed. Here, the folded RNA chain used as the template-primer to synthesize msDNA, becomes associated with a random nick in the chromosome, perhaps generated by a phage integrase or perhaps because of a defect in a repair system of the host. The exposed 3′-OH at the chromosome nick then serves as an alternate primer to intiate the synthesis of a partial copy of msDNA by the retron-encoded RT. The partial cDNA copy then becomes incorporated into the chromosome at the site of the nick.

Conclusions and speculations

There is now resounding evidence, from the DNA sequence analysis of eukaryotic genomes, that the process of reverse transcription has had a powerful impact on determining the structure, evolution, and gene content of the human genome as well as many other eukaryotic genomes. It seems especially relevant now to ask if reverse transcription has likewise been a force shaping the genomes of the prokaryotes. This is particularly interesting given the well documented presence of active RT encoding genetic elements in a wide diversity of bacterial genomes.

There are a number of examples of bacterial DNA sequences that have been postulated to be generated by reverse transcription. For example, many bacterial chromosomes contain numerous short, repeated DNA sequences. These include the REP-type sequences, which are 30-bp palindromic repeats, with greater than 500 copies per *E. coli* chromosome (Lupski and Weinstock, 1992). These repetitive DNAs are located in extragenic regions of the bacterial chromosome and have been proposed to be duplicated by a reverse transcription mechanism (Higgins et al., 1988). Another example is the mobile gene cassettes associated with integron elements in bacteria (Hall and Collis, 1995). These gene cassettes are usually DNAs encoding a gene, but no promoter, that mediates resistance to an antibiotic. Also present is a 59-base sequence that allows the integron element to "capture" and incorporate these gene cassettes by site-specific recombination. The formation of these gene cassettes has also been proposed to occur via reverse transcription of mRNA molecules (Recchia and Hall, 1997). However, little or no evidence exists to determine how these cassettes and repeat DNAs are produced in bacteria. Likewise, no source of RT activity responsible for the synthesis of these DNAs has been identified.

There are now a few documented examples of DNAs produced by reverse transcription in bacteria. This includes the demonstration of retrotransposition of a group II intron in *Lactococcus lactis* (Ichiyanagi et al, 2002). Here, a cDNA copy of the intron RNA is synthesized by the intron encoded protein (RT) during one of the multi-step events mobilizing transfer of the intron element to a new site. A second example appears to involve a phage-encoded RT that may generate mutational variations in a phage tail fiber protein (Liu et al., 2002). This allows the phage to phenotypically switch its ability to infect new variants of its bacterial host *Bordetella* and perhaps other bacterial hosts. However, retrons and the msDNAs that they produce are the first example of a bacterial DNA that has been shown both in vivo and in vitro to be synthesized by reverse transcription (Lampson et al., 1989; Shimamoto et al., 1995). The presence of numerous partial copies of the msDNA *(msd)* sequence, scattered about the chromosome of the myxobacterium *N. exedens*, are very likely retrotransposed DNAs. These truncated msDNA sequences may be an emerging family of repetitive DNAs that have been produced by reverse transcription. Thus, retron RT may be responsible for generating these partial msDNA sequences and perhaps other types of repetitive DNAs found in bacterial genomes. Although much better evidence is needed, these msDNA sequence copies may be duplicated in a manner similar to SINE elements found in eukaryotic genomes. Perhaps a new appreciation is needed for the impact of reverse transcription, a process that has had such a profound effect on the structure of most eukaryotic genomes, on prokaryotic genomes. RTs, such as those encoded by retrons, may play a role in spontaneous mutations, duplication of DNA sequences, and perhaps other changes to the bacterial genome. Perhaps msDNA, the product from retrons, may not be simply a selfish DNA, but may have a role in the cell. These are important questions to be answered for our full understanding of the role of retrons in bacteria.

References

Ahmed AM, Shimamoto T: msDNA-St85, a multicopy single-stranded DNA isolated from *Salmonella enterica* serovar Typhimurium LT2 with the genomic analysis of its retron. FEMS Microbiol Lett 224: 291–297 (2003).

Belfort M, Derbyshire V, Cousineau B, Lambowitz A: Mobile introns: pathways and proteins, in Craig NL, Craigie R, Gellert M, Lambowitz A (eds): Mobile DNA II (American Society for Microbiology Press, Washington DC 2002).

Boyd DA, Peters GA, Ng L-K, Mulvey MR: Partial characterization of a genomic island associated with the multidrug resistance region of *Salmonella enterica* Typhimurium DT104. FEMS Microbiol Lett 189:285–291 (2000).

Dai L, Zimmerly S: The dispersal of five group II introns among natural populations of *Escherichia coli*. RNA 8:1294–1307 (2002).

Dhundale A, Furuichi T, Inouye S, Inouye M: Distribution of multicopy single-stranded DNA among myxobacteria and related species. J Bacteriol 164: 914–917 (1985).

Dhundale A, Lampson B, Furuichi T, Inouye M, Inouye S: Structure of msDNA from *Myxococcus xanthus*: evidence for a long self-annealing RNA precursor for the covalently linked, branched RNA. Cell 51:1105–1112 (1987).

Dhundale A, Furuichi T, Inouye M, Inouye S: Mutations that affect production of branched RNA-linked msDNA in *Myxococcus xanthus*. J Bacteriol 170:5620–5624 (1988).

Dodd IB, Egan JB: The *Escherichia coli* retrons Ec67 and Ec86 replace DNA between the *cos* site and a transcription terminator of a 186-related prophage. Virology 219:115–124 (1996).

Eickbush TH: Repair by retrotransposition. Nature Genet 31:126–127 (2002).

Feng G, Tsui H-CT, Winkler ME: Depletion of the cellular amounts of the MutS and MutH methyl-directed mismatch repair proteins in stationary-phase *Escherichia coli* K-12 cells. J Bacteriol 178: 2388–2396 (1996).

Furuichi T, Dhundale A, Inouye M, Inouye S: Branched RNA covalently linked to the 5′ end of a single-stranded DNA in *Stigmatella aurantiaca*: structure of msDNA. Cell 48:47–53 (1987a).

Furuichi T, Inouye S, Inouye M: Biosynthesis and structure of stable branched RNA covalently linked to the 5′ end of multicopy single-stranded DNA of *Stigmatella aurantiaca*. Cell 48:55–62 (1987b).

Hall RM, Collis CM: Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination. Mol Microbiol 15:593–600 (1995).

Halling C, Calendar R: Bacteriophage P2 *ogr* and P2 δ gene act independently and are essential for P4 multiplication. J Bacteriol 172:3549–3558 (1990).

Herzer PJ: Starvation-induced expression of retron-Ec107 and the role of ppGpp in multicopy single-stranded DNA production. J Bacteriol 178:4438–4444 (1996).

Herzer PJ, Inouye S, Inouye M, Whittam TS: Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. J Bacteriol 172:6175–6181 (1990).

Herzer PJ, Inouye S, Inouye M: Retron-Ec107 is inserted into the *Escherichia coli* genome by replacing a palindromic 34 bp intergenic sequence. Mol Microbiol 6:345–354 (1992).

Higgins CF, McLaren RS, Newbury SF: Repetitive extragenic palindromic sequences, mRNA stability and gene expression: evolution by gene conversion? Gene 72:3–14 (1988).

Hsu M-Y, Inouye S, Inouye M: Structural requirements of the RNA precursor for the biosynthesis of the branched RNA-linked msDNA of *Myxococcus xanthus*. J Biol Chem 264:6214–6219 (1989).

Hsu M-Y, Inouye M, Inouye S: Retron for the 67-base multicopy single-stranded DNA from *Escherichia coli*: A potential transposable element encoding both reverse transcriptase and Dam methylase functions. Proc Natl Acad Sci USA 87:9454–9458 (1990).

Hsu M-Y, Xu C, Inouye M, Inouye S: Similarity between the *Myxococcus xanthus* and *Stigmatella aurantiaca* reverse transcriptase genes associated with multicopy, single-stranded DNA. J Bacteriol 174:2384–2387 (1992).

Ichiyanagi K, Beauregard A, Lawrence S, Smith D, Cousineau B, Belfort M: Retro-transposition of the Ll.LtrB group II intron proceeds predominantly via reverse splicing into DNA targets. Mol Microbiol 46:1259–1272 (2002).

Inouye S, Herzer PJ, Inouye M: Two independent retrons with highly diverse reverse transcriptases in *Myxococcus xanthus*. Proc Natl Acad Sci USA 87: 942–945 (1990).

Inouye S, Sunshine MG, Six EW, Inouye M: Retron-phage φR73: an *E. coli* phage that contains a retroelement and integrates into a tRNA gene. Science 252:969–971 (1991).

Inouye S, Hsu M-Y, Xu A, Inouye M: Highly specific recognition of primer RNA structures for 2′-OH priming reaction by bacterial reverse transcriptases. J Biol Chem 274:31236–31244 (1999).

Kawaguchi T, Herzer PJ, Inouye M, Inouye S: Sequence diversity of the 1.3 Kb retron (retron-Ec107) among three distinct phylogenetic groups of *Escherichia coli*. Mol Microbiol 6:355–361 (1992).

Kazazian HH: L1 retrotransposons shape the mammalian genome. Science 289:1152–1153 (2000).

Kim K, Jeong D, Lim D: A mutational study of the site-specific cleavage of Ec83, a multicopy single-stranded DNA (msDNA): nucleotides at the msDNA stem are important for its cleavage. J Bacteriol 179:6518–6521 (1997).

Lampson BC, Rice SA: Repetitive sequences found in the chromosome of the myxobacterium *Nannocystis exedens* are similar to msDNA: a possible retrotransposition event in bacteria. Mol Microbiol 23: 813–823 (1997).

Lampson BC, Inouye M, Inouye S: Reverse transcriptase with concomitant ribonuclease H activity in the cell-free synthesis of branched RNA-linked msDNA of *Myxococcus xanthus*. Cell 56:701–707 (1989a).

Lampson BC, Sun J, Hsu M-Y, Vallejo-Ramirez J, Inouye S, Inouye M: Reverse transcriptase in a clinical strain of *E. coli*: its requirements for production of branched RNA-linked msDNA. Science 243: 1033–1038 (1989b).

Lampson BC, Viswanathan M, Inouye M, Inouye S: Reverse transcriptase from *Escherichia coli* exists as a complex with msDNA and is able to synthesize double-stranded DNA. J Biol Chem 265:8490–8496 (1990).

Lampson BC, Inouye M, Inouye S: Survey of multicopy single-stranded DNAs and reverse transcriptase genes among natural isolates of *Myxococcus xanthus*. J Bacteriol 173:5363–5370 (1991).

Lampson B, Inouye M, Inouye S: The msDNAs of bacteria. Prog Nucl Acid Res Mol Biol 67:65–91 (2001).

Lampson BC, Xu C, Rice SA, Inouye S: A partial copy of msDNA from a new retron element is likely a retrotransposed DNA found in the myxobacterium *Nannocystis exedens*. Gene 299:251–261 (2002).

Lim D: Structure of two retrons of *Escherichia coli* and their common chromosome insertion site. Mol Microbiol 5:1863–1872 (1991).

Lim D: Structure and biosynthesis of unbranched multocopy single-stranded DNA by reverse transcriptase in a clinical *Escherichia coli* isolate. Mol Microbiol 6:3531–3542 (1992).

Lim D, Maas WK: Reverse transcriptase-dependent synthesis of a covalently linked, branched DNA-RNA compound in *E. coli* B. Cell 56:891–904 (1989).

Lima TMO, Lim D: Isolation and characterization of host mutants defective in msDNA synthesis: role of ribonuclease H in msDNA syntheisis. Plasmid 33:235–238 (1995).

Lima TMO, Lim D: A novel retron that produces RNA-less msDNA in *Escherichia coli* using reverse transcriptase. Plasmid 38:25–33 (1997).

Liu M, Devra R, Doulator SR, Gingery M, Eiserling FA, Preston A, Maskell DJ, Simons RW, Cotter PA, Parkhill J, Miller JE: Reverse transcriptase-mediated tropism switching in *Bordetella* bacteriophage. Science 295:2091–2094 (2002).

Lupski SR, Weinstock GM: Short, interspersed repetitive DNA sequences in prokaryotic genomes. J Bacteriol 174:4525–4529 (1992).

Maas WK, Wang C, Lima T, Zubay G, Lim D: Multicopy single-stranded DNAs with mismatched base pairs are mutagenic in *Escherichia coli*. Mol Microbiol 14:437–441 (1994).

Maas WK, Wang C, Lima T, Hach A, Lim D: Multicopy single-stranded DNA of *Escherichia coli* enhances mutation and recombination frequencies by titrating MutS protein. Mol Microbiol 19:505–509 (1996).

Mao J-R, Inouye S, Inouye M: Enhancement of frameshift mutation by the over-production of msDNA in *Escherichia coli*. FEMS Microbiol Lett 144:109–115 (1996).

Mao J-R, Inouye S, Inouye M: msDNA-Ec48, the smallest multicopy single-stranded DNA from *Escherichia coli*. J Bacteriol 179:7865–7868 (1997).

Martinez-Abarca F, Toro N: Group II introns in the bacterial world. Mol Microbiol 38:917–926 (2000).

Matiasovicova J, Faldynova M, Pravcova M, Darpiskova R, Kolackova I, Damborsky J, Rychlik I: Retron reverse transcriptase *rrtT* is ubiquitous in strains of *Salmonella enterica* serovar Typhimurium. FEMS Microbiol Lett 223:281–286 (2003).

Ogata H, Audic S, Barbe V, Artigueuave J, Fournier P-E, Raoult D, Claverie JM: Selfish DNA in protein-coding genes of *Rickettsia*. Science 290:347–350 (2000).

Recchia GD, Hall RM: Origins of the mobile gene cassettes found in integrons. Trends Microbiol 5:389–394 (1997).

Rest JS, Mindell DP: Retroids in Archaea: phylogeny and laterial origins. Mol Biol Evol 20:1134–1142 (2003).

Rice SA, Lampson BC: Phylogenetic comparison of retron elements among the myxobacteria: evidence for vertical inheritance. J Bacteriol 177:37–45 (1995).

Rice SA, Bieber J, Chun J-Y, Stacey G, Lampson BC: Diversity of retron elements in a population of rhizobia and other Gram-negative bacteria. J Bacteriol 175:4250–4254 (1993).

Rychlik I, Sebkova A, Gregorova D, Karpiskova R: Low-molecular-weight plasmid of *Salmonella enterica* serovar Enteritidis codes for retron reverse transcriptase and influences phage resistance. J Bacteriol 183:2852–2858 (2001).

Sawyer SA, Dykhuizen DE, DuBose RF, Green L, Mutangadura-Mhlanga T, Wolczyk DF, Hartl DL: Distribution and abundance of insertion sequences among natural isolates of *Escherichia coli*. Genetics 115:51–63 (1987).

Selander RK, Caugant DA, Whittam TS: *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology (American Society for Microbiology Press, Washington DC 1987).

Shimamoto T, Shimada M, Inouye M, Inouye S: The role of ribonuclease H in multicopy single-stranded DNA synthesis in retron-Ec73 and retron-Ec107 of *Escherichia coli*. J Bacteriol 177:264–267 (1995a).

Shimamoto T, Inouye M, Inouye S: The formation of the 2′,5′-phosphodiester linkage in the cDNA priming reaction by bacterial reverse transcriptase in a cell-free system. J Biol Chem 270:581–588 (1995b).

Shimamoto T, Kobayashi M, Tsuchiya T, Shinoda S, Kawakami H, Inouye S, Inouye M: A retroelement in *Vibrio cholerae*. Mol Microbiol. 34:631–632 (1999).

Steitz TA: DNA polymerases: structural diversity and common mechanisms. J Biol Chem 274:17395–17397 (1999).

Sun J, Inouye M, Inouye S: Association of a retroelement with a P4-like cryptic prophage (retronphage φR73) integrated into the selenocystyl tRNA gene of *Escherichia coli*. J Bacteriol 173:4171–4181 (1991).

Temin HM: Retrons in bacteria. Nature 339:254–255 (1989).

Weiner AM: SINEs and LINEs: The art of biting the hand that feeds you. Curr Opin Cell Biol 14:343–350 (2002).

Woese CR: Bacterial evolution. Microbiol Rev 51:221–271 (1987).

Xiong Y, Eickbush TH: Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J 9:3353–3362 (1990).

Yamanaka K, Shimamoto T, Inouye S, Inouye M: Retrons, in Craig NL, Craigie R, Gellert M, Lambowitz AM (eds): Mobile DNA II, pp 784–795 (American Society for Microbiology Press, Washington DC 2002).

Yee T, Furuichi T, Inouye S, Inouye M: Multicopy single-stranded DNA isolated from a Gram-negative bacterium, *Myxococcus xanthus*. Cell 38:203–209 (1984).