**Glossary**

**Group I intron** Self-splicing ribozymes that can catalyze their own excision from mRNA, tRNA or rRNA precursor transcripts. Require an external GTP to initiate the splicing reaction and the assistance of protein factors are required for efficient in vivo splicing.

**Group II intron** Self-splicing ribozymes that mechanistically resemble splicesome introns forming a lariat RNA molecule with an Adenine residue branch point. Splicing does not require GTP but protein factors are required for rapid splicing in vivo.

**Homing endonucleases** Promote mobility by inducing DNA breaks in intronless alleles thus stimulating mobility by homologous recombination based mechanisms. Encoded by group I, II and archeal introns, also components of inteins and can be expressed from freestanding genes.

**Introns** A sequence that is present within a gene but removed during RNA processing of the gene’s transcript.

**Maturases** Group I and II intron encoded proteins that assist the intron RNA in obtaining the proper configuration to be splicing competent.

**Ribozymes** Catalytic RNA molecules first discovered in 1981 by Thomas Cech who found a group I intron RNA molecules that was capable of catalyzing RNA cleavage and splicing reaction in sequence-specific ways.

**Retrotransposition** Mobile elements (such as group II introns, Long and Short interspersed repetitive elements such as LINES and SINES, retrotransposons, retroviruses) that move into new sites by involving an RNA intermediate that has to be integrated back into the genome with the aid of reverse transcriptase.

**Targetrons** They are commercially available, site specific gene knockout systems based on genetically engineered group II introns that are programmed to retrohome into specific sites (genes).

**Twintron** Introns inserted within introns can be named twintrons, but a more strict definition for twintron implies that splicing of the external member of the twintron can only proceed after splicing of the internal member.

---

**Overview of Mobile Introns**

Group I and group II introns are commonly referred to as mobile introns as they can move from an intron containing allele to a cognate allele that lacks the intron. The mobility mechanisms are usually site specific, thus the term “homing.” and mobility is mediated by proteins encoded within the mobile introns. Intron homing is different from intron transposition, a term that is applied when an intron inserts itself at a different site; this is also referred to as ectopic integration. Ecotopic integration events are viewed to be important for these elements in order to persist by occupying new sites and thus avoiding extinction as they are viewed to be neutral elements that quickly accumulate mutations due to lack of selection. Ecotopic integration events can also generate twintron-configurations where two ribozyme type introns co-exist at the same locus.

Group I and II introns can be distinguished from each other by their sequences, secondary and tertiary RNA structures, and splicing mechanisms. These introns are potential ribozymes catalyzing their own removal from the precursor transcripts. Mobile introns contain two functional components the autocatalytic self-splicing intron RNA and an intron encoded protein (IEP), such as homing endonucleases (HEs) or in group II introns a complex multifunctional reverse transcriptase-like protein (RT).

The current view is that pre-existing group I introns recruited ORFs that encoded HEs providing the intron with a mechanism for mobility. However, once an intron and its ORF have been established within a specific host gene, selection might favor the development of maturase activity over endonuclease activity because correct and efficient splicing would lessen the impact of the intron on the host gene. Maturases are thought to facilitate splicing by promoting proper folding of the intron RNA. Group I introns typically move by a DNA based mobility mechanism that involves homologous recombination, whereas the movement of group II introns involves an RNA intermediate and homologous recombination is not involved.

Mobile introns have been noted in Eubacteria and their phages and within the Archaea. In Eubacteria group II introns frequently are encountered within other mobile elements such as conjugative transposons. With the exception of nuclear rDNA group I introns in some protozoans and fungi, group I and II introns in eukaryotes tend to be restricted to organelar genomes, although these introns are rare among the metazoan. Mobile introns contribute toward genomic variability and they may

---

intron-less cognate allele, and the formation between the intron lariat and the intron-encoded protein (possibly host factors. During expression of the host gene, the intronic ORF is translated and a ribonucleoprotein particle (RNP) is formed within numerous group I introns, and LAGLIDADG HEs and the H-N-H domain are found within numerous group I introns, and LAGLIDADG HEs and the H-N-H domain are present within the ORFs of some group II introns. Additional HE-like proteins have been described, the PD-(D/E)XK HEs are found in bacterial tRNA group I introns, the very-short patch repair (Vsr) endonucleases (a predicted family of phage HEs based on metagenomic), and the Holliday junction resolvase-like HEs found in some phage introns.

The ability of HEGs to move independently of their ribozyme counterparts to form new composite mobile units along with some allowance for degeneracy at their DNA target sites provides the flexibility needed for HEGs and mobile introns to invade new sites. Recent work shows that there are complex interactions between the expression of IEPs, intron splicing, the expression of the host gene and cellular metabolism suggesting a mutualistic relationship between the host genome and the mobile intron.

**Group II Introns: Retro-Homing**

Group II introns in general have conserved secondary structures at the RNA level, that can be visualized as six stem-loop domains (domains I to VI) emerging from a central wheel. The RT ORFs if present tend to be embedded within domain IV but in a few instances they have been observed in domain II. There are variants of group II introns that lost domains II to V and these typically range in size from 93 to 118 nt but they splice in a similar manner as standard group II introns. Mobile group II introns typically encode a multifunctional protein which contains a segment homologous to reverse transcriptases (RT), domain X which has been implicated in maturase activity and the En domain which contains a potential zinc finger which has endonuclease activity. However, the En domain is absent in some fungal group II introns.

Both splicing and mobility of group II introns require the catalytic activity of the intron RNA, the intron-encoded protein, and possibly host factors. During expression of the host gene, the intronic ORF is translated and a ribonucleoprotein particle (RNP) is formed between the intron lariat and the intron-encoded protein (Fig. 1b). The RNP recognizes a target homing site, typically an intron-less cognate allele, and the first cut is made by the 3' end of the intron RNA. This initiates a reverse splicing reaction.
whereby the intron RNA is inserted into the sense DNA strand. The En domain cleaves the antisense DNA strand, generating a free 3′-OH that serves as a primer for the RT. Eventually, the host DNA repair machinery will remove the RNA and fill in any gaps.

Intron-catalyzed splicing proceeds by two transesterification reactions and leads to the excision of the intron as a branched, or lariat, molecule with a characteristic 2′-5′ phosphodiester bond, as in the branched pathway, or in a linear form in the hydrolytic pathway. A requirement for the group II intron splicing reaction is that the 5′ and 3′ exon sequences flanking the splice site are bound at the active site by base pair interactions involving sequence elements embedded within domain I of the intron RNA. These exon binding sites (EBS 1 and 2) bind to the corresponding intron binding sites (IBS1 and 2) located within the 5′ exon directly upstream of the 5′ intron/exon junction. These base-pairing interactions are required for splicing, reverse splicing into RNA and for insertion into DNA target sites during retrohoming.
<table>
<thead>
<tr>
<th>Structural family</th>
<th>Enzyme</th>
<th>Cutting site and products of the cut</th>
<th>Subcellular location</th>
<th>Source</th>
<th>Biological domain</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAGLIDADG</td>
<td>I-CreI</td>
<td>5’ - - -CTGGGTTCAAAACGTCGTGA GACAGTTTGG- - - 3’</td>
<td>Chloroplast</td>
<td>Chlamydomonas reinhardtii</td>
<td>Eukaryota</td>
<td>Jurica et al. (1998)</td>
</tr>
<tr>
<td></td>
<td>I-AniI</td>
<td>5’ - - -TTGAGGAGGTTCGCTGTAATAA- - - 3’</td>
<td>Mitochondria</td>
<td>Aspergillus nidulans</td>
<td>Eukaryota</td>
<td>Bolduc et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>I-DmoI</td>
<td>5’ - - -ATGCCTTGGCGCTAAGGTAAAATTA- - - 3’</td>
<td>Chromosome</td>
<td>Desulfurococcus mobilis</td>
<td>Archaea</td>
<td>Silva and Belfort (2004)</td>
</tr>
<tr>
<td></td>
<td>PI-Pfl</td>
<td>5’ - - -GAAGATGGGAGGACGCCATCAAC- - - 3’</td>
<td>Chromosome (Intein)</td>
<td>Pyrococcus furiosus</td>
<td>Archaea</td>
<td>Komori et al. (1999)</td>
</tr>
<tr>
<td>GIY-YIG</td>
<td>I-TevI</td>
<td>5’ - - -AGGAGTATCAACGCTGAGTCAGTGCA- - - 3’</td>
<td>Phage</td>
<td>Escherichia coli</td>
<td>Bacteria</td>
<td>Vanroey et al. (2002)</td>
</tr>
<tr>
<td></td>
<td>I-Hm1l</td>
<td>5’ - - -TCACCATTGCTGAGTCAGGTTGCAA- - - 3’</td>
<td>Phage</td>
<td>Bacillus subtilis</td>
<td>Bacteria</td>
<td>Shen et al. (2004)</td>
</tr>
<tr>
<td></td>
<td>I-SceV</td>
<td>5’ - - -TCACTTGGCGCTAAGGTAAAATTA- - - 3’</td>
<td>Mitochondria</td>
<td>Saccharomyces cerevisiae</td>
<td>Eukaryota</td>
<td>Liang et al. (1996)</td>
</tr>
<tr>
<td>His-Cys Box</td>
<td>I-PpoI</td>
<td>5’ - - -TATATACAAAGAGGATCTTTAAGGATCGCAAT- - - 3’</td>
<td>Nuclear</td>
<td>Physarum polycephalum</td>
<td>Eukaryota</td>
<td>Wittmayer et al. (1998)</td>
</tr>
<tr>
<td>Artificial (LAGLIDADG)</td>
<td>H-DreI</td>
<td>5’ - - -AAACACGCTGCTCATCGCGGTTTAA- - - 3’</td>
<td>N/A</td>
<td>Escherichia coli</td>
<td>Bacteria</td>
<td>Chevalier et al. (2002)</td>
</tr>
</tbody>
</table>
Intron Mobility by Reverse Splicing and Retrotransposition

Group I introns may also transpose into new sites (ectopic integration) involving RNA intermediates by reverse splicing (Fig. 2a). Reverse splicing allows a free group I intron RNA to insert into a homologous or heterologous RNA; this mode of mobility requires complementary base pairing between the intron and the exon RNA sequences. This mechanism requires the additional steps of reverse transcription of the RNA and the integration of the cDNA into the genome by a recombination step that replaces the “intron-less copy” with the intron-plus cDNA. As reverse splicing requires less homology (4–6 nucleotides) this mechanism allows introns to spread more easily into heterologous sites.

Group II intron RNAs have been shown to retrotranspose by reverse splicing into RNA molecules (Fig. 2b), a mechanism that requires less specificity at the target site and thus allows for retrotransposition of introns to new sites within the genome. Group II introns can also reverse splice into single stranded DNA (at the replication fork; Fig. 2c) facilitation their dispersal into new sites and transmission to the next generation.

**Fig. 2** (a) and (b) Reverse splicing of group I and II intron RNAs (red lines) into RNA substrates (gray and black wavy lines). Target-site recognition requires the interaction of the internal guide sequence (IGS) of the group I intron RNA with the recipient RNA or the exon binding sequence (EBS) of the group II intron RNA with the intron binding sequence (IBS) in the recipient RNA molecule. Reverse splicing requires a reverse transcription step to generate cDNA followed by homologous recombination that integrates the intron containing cDNA into the genome. (c) Group II intron RNA can also reverse splice into single-stranded DNA (black solid lines) as found at the DNA replication fork and again a reverse transcriptase step is required for cDNA synthesis but there is no need for endonuclease (En) activity and for homologous recombination in order for the integration of the intron to be completed (black and red). Reverse transcription is primed by the nascent leading strand of the DNA replication fork (black-red dashed line). Mechanisms depicted in (a)–(c) have a lower homology requirement for target-site recognition and therefore can promote the retrotransposition of introns into new locations (ectopic integration) within the genome.
Applications and Biotechnology

Mutations within group II intron EBS motifs allow for changing intron target specificities. Engineered group II introns, termed “targetron,” have been developed, which allow for gene targeted mutagenesis in a variety of bacteria. The targetron is introduced into a bacterial cell by means of a compatible plasmid vector and the group II intron has been programmed to insert into a specific target site/gene. Targetron-like systems are being developed as gene delivery systems, whereby genes are incorporated into domain IV of the intron and the intron encoded protein (RT) has been relocated either on a second vector or into a different position within the same vector. If successful these types of systems would allow for site specific DNA insertions and provide new tools in genetically manipulating bacteria and eukaryotes. Gene replacement strategies are also being developed whereby a modified RT-deficient group II intron is used to introduce targeted site specific double-stranded breaks. These breaks will induce the host systems DNA DSB repair system involving homologous recombination. A co-transformed DNA fragment can be engineered to be the template for homologous recombination and thus replacing the “damaged” segment.

Homing endonucleases require long DNA recognition sites and therefore cut infrequently within a genome; this makes them useful for DNA engineering and genomics. HEs can be engineered to cleave at desired locations and therefore HEs can become site specific tools that can be used to target specific genes for mutagenesis. Gene replacement strategies are also being developed that would allow for therapeutic applications for HEs that target genes associated with human diseases.

References


Further Reading


Relevant Websites

http://www.tf.ucalgary.ca/group2introns/
Mobile Group II Intron Database.
http://www.mcb.utexas.edu/
The University of Texas at Austin – Comparative RNA Web Site and Project.
http://www.wuhan.edu.cn/gissd/help.html
Wuhan University, China – Group I Intron Sequence and Structure Database.