RNA Editing

The Discovery of RNA Editing Brought into Question the Central Dogma of Molecular Genetics

One of the basic tenets of molecular genetics is that the nucleotide sequence in the messenger RNA should be a perfect copy of the sequence in the DNA, as determined by the rules of base pairing. The first challenge to this idea came with the discovery of intervening sequences within genes of higher organisms which are precisely spliced out of the mRNA, and the coding RNA fragments, or exons, are then joined together to create the complete gene. This discovery did not really contradict the idea that the mRNA sequence is a copy of the DNA sequence since the exons do represent perfect copies.

However, evidence obtained in the past decade with an ancient group of parasitic flagellated protozoa, the kinetoplastids, indicates that the sequence of nucleotides in mRNAs in coding regions can be modified after transcription in quite dramatic ways, a phenomenon known as RNA editing. These results initially led some people to question the central dogma of molecular genetics, which states that genetic information flows from DNA \( \rightarrow \) RNA \( \rightarrow \) protein, since the genetic information for the sequence changes in the mRNAs and hence the sequence of amino acids in the proteins did not appear to be encoded in the nucleic acids of the organism. These results can now be explained in terms of a model which is consistent with the central dogma but yet has some novel and interesting features.

Other examples of nonencoded sequence modifications...
of RNAs in other organisms were also reported and were also termed RNA editing, although the mechanisms appear to be quite distinct. In some cases, the site specificity of the editing events can be explained in terms of proteins that recognize specific RNA sequences or RNA structures, but in other cases the site specificity is a mystery.

A Personal History of the Investigation of RNA Editing in Trypanosomes

Trypanosomes

Trypanosomes belong to a large group of parasitic protozoa known as kinetoplastid protozoa. The term trypanosome or trypanosomatid is actually the name for cells belonging to the genus *Trypanosoma* in this family, but it is loosely used to describe any kinetoplastid protozoan. The name kinetoplastid was derived from the fact that early investigators saw a small granule or kinetoplast at the base of the flagellum when they stained cells with certain dyes. The kinetoplast actually represents a portion of the single complex mitochondrion of the cell that contains a huge compact mass of mitochondrial DNA, which is what stains with dyes.

Trypanosomes are both repelling and fascinating. Several species are dangerous human parasites, causing such diseases as African sleeping sickness, Chagas’s disease, and leishmaniasis, a skin ulceration that in some cases can metastasize to the mouth and nose and cause considerable damage. There is no vaccine for any of these diseases and the treatments, when they exist, are not very satisfactory. In Africa, domestic cattle (but not wild animals) succumb to a trypanosome species transmitted by the tsetse fly, and even coconut palms are affected by one type of trypanosome, *Phytomonas*. In fact, most of the early work on trypanosomes was done by medical doctors interested in the diseases. However, many trypanosome species are only parasites of insects and have no medical significance. Others are parasites of both insects and vertebrates.

Trypanosomes are fascinating to the modern cellular and molecular biologist since they represent one of the most primitive lines of descent of eukaryotic cells. Perhaps for this reason, they possess many truly unusual features not found in other eukaryotic cells. For example, the trypanosomes have only a single mitochondrion which contains a truly unique type of mitochondrial DNA known as kinetoplast DNA. All modern-day eukaryotic unicellular organisms whose ancestors arose prior to the ancestors of trypanosomes, such as *Giardia* and *Trichomonas*, lack mitochondria and therefore do not respire oxygen. It is possible that the lineage which gave rise to the Euglenozoa flagellated protozoa (which includes the kinetoplastid protozoa and also *Euglena*, a green flagellated protozoan) approximately 1 billion years ago was very closely related to that primitive ancestral cell which engulfed a respiring bacterial cell to create the obligate intracellular organelle now known as the mitochondrion.

Kinetoplast DNA

I became interested in the unusual mitochondrion of trypanosomes as a graduate student. At that time, all that was known was that this mitochondrion contained a large compact mass of DNA known as kinetoplast DNA that stained well with certain dyes. Our model system was the nonpathogenic trypanosome, *Leishmania tarentolae*, which was originally a parasite of a gecko. We and others soon showed that this DNA consists of thousands (5000–10,000) of small circular DNA molecules known as minicircles, all linked together by catenation-like rings in a chain forming a giant network of DNA, and a smaller number (20–50) of larger circular DNA molecules known as maxicircles, also linked to the network (Simpson and da Silva, 1971) (Fig. 1). The minicircles, although identical in size in any one species, consist of molecules having different sequences even in the same network. Minicircles range in size from 460 to 2500 bp in different species, and each minicircle is organized into one or more conserved regions and variable regions. A region of the minicircle is inherently bent (Marini et al., 1982), the genetic role of which is still unclear. Minicircles did not appear to encode information for proteins but did replicate very well and have been keeping several labs busy for years trying to elucidate the mechanisms of replication and
segregation of the network (Ryan et al., 1988; Shapiro and Englund, 1995). The genetic function of minicircles proved elusive and was not solved until the discovery of guide RNAs in 1990 (Blum et al., 1990).

The maxicircle molecule, which ranges in size from 23,000 to 36,000 bp in different species, appears to represent the informational DNA molecule in the mitochondrion and contains several of the same genes found in a human mitochondrial DNA molecule: the large and small mitochondrial ribosomal RNAs, three subunits of cytochrome oxidase, cytochrome b, several subunits of NADH dehydrogenase, a subunit of the F1–F0 ATPase, and several unidentified proteins. All of the identified structural genes are involved with electron transport and oxidative phosphorylation in the inner membrane of the organelle, as in human cells.

**First Indications of Cryptogenes**

It was somewhat disturbing in the mid-1980s that several genes found in human and yeast mitochondrial genomes were apparently not present in trypanosome mitochondria, such as the mitochondrial tRNA genes. The apparent absence of tRNA genes was surprising since all mitochondrial studied until that time contained tRNAs which were involved in mitochondrial protein synthesis. To examine whether some of the minicircle molecules or regions of the maxicircle which we had not yet sequenced encoded whether some of the minicircle molecules or regions of involvement in mitochondrial protein synthesis. To examine whether some of the minicircle molecules or regions of the maxicircle which we had not yet sequenced encoded tRNAs, we hybridized labeled minicircle and maxicircle the maxicircle which we had not yet sequenced encoded whether some of the minicircle molecules or regions of involvement in mitochondrial protein synthesis. To examine whether some of the minicircle molecules or regions of involvement in mitochondrial protein synthesis. To examine whether some of the minicircle molecules or regions of involvement in mitochondrial protein synthesis. To examine whether some of the minicircle molecules or regions of involvement in mitochondrial protein synthesis. To examine whether some of the minicircle molecules or regions of involvement in mitochondrial protein synthesis. To examine whether some of the minicircle molecules or regions of involvement in mitochondrial protein synthesis. To examine whether some of the minicircle molecules or regions of involvement in mitochondrial protein synthesis. To examine whether some of the minicircle molecules or regions of involvement in mitochondrial protein synthesis.
that the process occurred after transcription since transcription moves in a 5' to 3' direction, and it explained why partially edited messengers could not be translated — the ribosome binding site was the last sequence to be created by editing. We found the same phenomenon to occur in the 5'-edited genes in *L. tarentolae* (Sturm and Simpson, 1990a). More than 400 partially edited mRNAs from the Cyb and COIII genes were sequenced that were generated by PCR amplification from mitochondrial RNA, and it was found that the region between the fully edited 3' side and the unedited 5' side of the mRNA, which we term the junction region, contained a variety of partially edited sequence patterns. In the case of the Cyb RNAs, almost all these patterns could be arranged into a precise 3' to 5' progression of editing (Fig. 3).

However, in the case of COIII, only 58% of the patterns showed this precise polarity. The rest showed unexpected editing patterns, in which U's were added at sites not edited in the mature transcript or U's were added in the 5' region before U's were added in the 3' region. Decker and Sollner-Webb (1990) analyzed partially edited mRNAs for Cyb and COIII from *T. brucei* and found a high percentage of unexpected patterns in the junction regions for both genes.

The realization that editing occurred after transcription and was both polarized and progressive suggested that a cut-and-splice mechanism was at work. However, the major problem was that there appeared to be no nucleic acid template for this newly added sequence information. It seemed as if sequence information was coming from nowhere and, in the case of pan-edited genes, entire new genes were being constructed. This clearly had profound implications for the central dogma of genetic information transfer.

**Discovery of Guide RNAs**

We had not given up on the central dogma, which had proven so resilient in the past. A simple computer search...
RNA editing

was performed of the known *L. tarentolae* maxicircle that was looking for short DNA sequences that could give rise to RNAs which could hybridize with either entire or portions of known edited RNA sequences. In addition to the classical Watson–Crick base pairs C–G and A–U, we decided to allow for G–U base pairs since these are bona fide base pairs in rRNAs and tRNAs, and this turned out to be the key (Fig. 4).

This computer search immediately found seven short sequences for four of the known edited genes scattered throughout the maxicircle between known genes (Blum et al., 1990). We soon had definitive evidence for the existence of small RNAs in the mitochondrion which were transcripts of these sequences. These RNAs also contained sequences at their 5′ end which could form duplex regions with the mRNAs just downstream of the preedited regions. We termed these sequences the “anchor regions” since these provided an ideal way to anchor the gRNA to

FIGURE 3  (a) Junction regions of mitochondrial RNAs for the gene Cyb, preedited (top) and completely edited (bottom). The sequences were obtained by RT-PCR of *L. tarentolae* kinetoplast RNA using a 3′ edited primer and a 5′ unedited primer. (b) Expected 3′ to 5′ editing. The different editing patterns (clones) are numbered in editing sequence order with the frequencies of patterns indicated on the right (in green). (c) Unexpected editing patterns (reproduced with permission from Sturm and Simpson, 1990a).

FIGURE 4  Alignment of the edited sequence of ND7 mRNA with the maxicircle sequence. (Top) The correspondence is shown, as originally observed, between the maxicircle sequence and a test DNA sequence. Transition mismatches are indicated in boxes. (Bottom) The perfect match that can be obtained if the fully edited mRNA (red) is aligned with the complementary RNA sequence of the maxicircle DNA (pink) and G–U base pairs are allowed (rectangles).
the mRNA by forming a double-stranded hybrid just downstream of the region that was to be edited.

These small RNAs had an unusual mobility in gel electrophoresis; they migrated in the form of 20–30 bands, each differing by a single nucleotide in size. We immediately realized that these were the same RNAs that had contaminated our tRNA preparations 2 years before (Simpson et al., 1989). The unusual mobility was soon found to be due to the presence of nonencoded oligo-[U] tails up to 24 nucleotides in length at the 3′ ends of the RNA molecules (Blum and Simpson, 1990). We called these molecules guide RNAs or gRNAs since they contained the sequence information for editing. Guide RNAs had come to the rescue of the central dogma! We were of course very excited, but at the same time we were somewhat chagrined that the answer to the secret of editing was not something completely new and earthshaking but rather something that obeyed the simple rules of base pairing. However, we still had to explain how these gRNAs could edit mRNA molecules.

At this time, we had identified seven gRNA genes scattered all over the maxicircle with no relation to the cryptogenes for which they encoded information. These gRNAs had information for four of the five known cryptogenes, but we could not find a gRNA for the 5′ edited COIII gene in *L. tarentolae*. Armed with the knowledge that minicircle DNA showed hybridization to small transcripts with the unusual gel mobility of gRNAs, the known minicircle sequences were searched and a gRNA gene was found in the D12 minicircle that contained sequence information for the first eight editing sites of the COIII mRNA (Sturm and Simpson, 1990b) (Fig. 5).

This was the first indication of a genetic function of the minicircle DNA and explained neatly the observed sequence heterogeneity of minicircle DNA: Each sequence class encoded a different gRNA within the variable region. A total of 17 different minicircle sequence classes of different gRNAs were located in the variable region precisely between three sets of 18 nucleotide inverted repeats, which are not present in *L. tarentolae* (Fig. 6). Another difference is that previous workers had shown that there were more than 300 different minicircle sequence classes in *T. brucei* versus the limited number found in *L. tarentolae.* This means that the total number of different gRNAs in *T. brucei* may be more than 900. One characteristic of the gRNAs in *T. brucei* is that they exhibit a large amount of redundancy. Redundant gRNAs have different sequences but contain the same editing information as a result of allowing G–U base pairs. A high percentage of redundant gRNAs have also been observed in *T. cruzi* (Avila and Simpson, 1995).

**Loss of Editing during Culture by Loss of Minicircles**

In the initial comparison of the mitochondrial genomes of *L. tarentolae* and *T. brucei*, we noted that there were several stretches of sequences that were relatively rich in G residues (Simpson et al., 1987). Three of these in *T. brucei* were determined to be the three hidden pan-edited cryptogenes, ND7, COIII, and MURF4, but there were an additional six G-rich regions that were located between known genes in both species. The transcript of G-rich region 6 in *L. tarentolae* is pan-edited by the addition of 117 U’s at 49 sites and the deletion of 32 U’s at 13 sites in three editing domains, producing an mRNA which encodes a protein for the small subunit of the mitochondrial ribosome (Maslov et al., 1992). The remaining G-rich regions, G1–G5, were shown to also be pan-edited cryptogenes in *T. brucei*, but no fully edited transcripts of these genes could be found in the old University of California (UC) lab strain of *L. tarentolae*. The mature edited mRNAs from G1, G2, and G5 in *T. brucei* proved to encode subunits of NADH dehydrogenase (Bhat et al., 1990; Souza et al., 1992, 1993; Read et al., 1992).

Examination of the recently isolated LEM125 strain of *L. tarentolae* suggested that the UC strain was genetically defective in editing of transcripts of the G1–G5 cryptogenes (Thiemann et al., 1994). The LEM125 strain contained fully edited G1–G5 mRNAs and in addition contained a more complex repertoire of minicircle-encoded gRNAs. At least 32 additional minicircle-encoded gRNAs were detected in the LEM125 strain which were absent in the UC strain (Tables 1 and 2). We hypothesized that specific minicircle sequence classes encoding gRNAs for the editing of the G1–G5 transcripts were lost during the long culture history of the UC strain, probably due to a lack of a requirement for the protein products during the culture stage of the *Leishmania* life cycle.

**Models for Mechanism of U-Insertion/Deletion Editing**

Prior to the discovery of gRNAs, we had isolated an enzyme activity from purified mitochondria of *L. tarentolae* which could add U’s to the 3′ terminus of any RNA molecule—a terminal uridylyl transferase or TUTase (Bakalara et al., 1989). This enzyme presumably was responsible for the addition of U’s to the 3′ end of the gRNAs. We also showed the presence of a mitochondrial RNA ligase which could covalently link together two RNA molecules and found that if the TUTase in the crude mitochondrial extract was inhibited by heparin or destroyed by digestion with proteinase K, a cryptic ribonuclease activity was activated.
RNA editing

**FIGURE 5** Minicircle-encoded gRNAs involved in the editing sites 1–8 of COIII mRNA. (a, top) A schematic representation of the D-12-1 minicircle of *L. tarentolae*. The polarity of the gRNA COIII gene (long red arrow) is indicated with respect to the conserved dodecamer CSB-3 sequence (shown in the rectangle), an origin of replication for one DNA strand (short red arrow). The blue band on the minicircle indicates the complete conserved sequence. The highlighted bar at the bottom indicates another sequenced region of the minicircle which permitted the determination of the polarity of the gene. (a, bottom) Complete sequence of the 61- to 240-bp region of the minicircle DNA including the possible gRNA sequence coding for the editing sites 1–8 of COIII along the D-12-1 minicircle. The red arrows indicate the 3' and 5' ends of the gRNA gene. (b) Edited COIII mRNA sequence of *L. tarentolae* (red) aligned with maxicircle DNA COIII sequence (blue) and of the gRNA (pink), showing the 17 editing sites and the inserted uridines, indicated as u. Deletions are indicated by asterisks (reproduced with permission from Sturm and Simpson, 1990b).

**FIGURE 6** Diagrams of genomic organization of minicircles from *L. tarentolae* and *T. brucei*. The locations of the conserved regions and the polarity of the CSB-3 sequence are shown together with the adjacent bent DNA region. The gRNA genes are indicated by orange arrows. In the *T. brucei* minicircle, the gRNA genes are enclosed by imperfect 18mer inverted repeats (gray arrows).
which cleaved the Cyb preedited mRNA within the pre-edited region (Simpson et al., 1992). A similar endonuclease activity was detected in mitochondrial extracts from T. brucei (Harris et al., 1992).

Armed with the knowledge of the existence of these enzymatic activities in the mitochondrion and the 3'→5' progression of editing on the mRNA, a model for the role of gRNAs in RNA editing was suggested (Blum et al., 1990). We called this the “enzyme cascade” model since it postulates a series of enzymatic reactions occurring in a multi-enzyme complex bound to the mRNA. We proposed that the initial interaction involves the formation of an anchor hybrid by the gRNA just 3' of the preedited region on the mRNA. In addition to RNA–RNA interactions involved in the formation of an anchor, we believe that protein factors which were found to be complexed to the gRNAs (Byrne et al., 1995; Bringaud et al., 1995) assist in this initial specific interaction, perhaps by recognizing secondary structures formed by the mRNA or by the gRNA–mRNA hybrid. The next step was proposed to be a specific cleavage at the first mismatched base in the mRNA which liberates a free 3' OH group. This cleaved mRNA fragment would be a substrate...
for the TUTase enzyme which could add one or more U’s to the 3’ end. These added U’s would then base pair with the guide A or G nucleotides in the gRNA, and then the two ends of the mRNA would be relegate by the RNA ligase. This would result in a zippering up of the double helix in a 3’ to 5’ direction (on the mRNA), and the whole process would then reinitiate at the next mismatched base (Fig. 7).

This model provided an explanation for the 3’ to 5’ polarity of pan-editing. It indicated that this polarity was due to the creation by the downstream gRNA of an edited mRNA sequence that was complementary to the anchor sequence of the adjacent upstream gRNA (Fig. 8). The model also explained the presence of unexpected editing patterns within the junction regions of partially edited mRNAs. We suggested that these patterns represented normal editing by inappropriate gRNAs or appropriate gRNAs in the wrong location or wrong reading frame — a process which we termed misediting and misguiding, and which is enhanced by the presence of “wobble” G–U and perhaps A–C base pairs (Sturm et al., 1992). The formation of an anchor hybrid by the wrong gRNA or the formation of a secondary anchor in the wrong location by the correct gRNA could lead to the formation of an unexpected editing pattern, which would terminate the editing process since a correct anchor for the next gRNA would not be formed. However, misedited sequences within the junction region could be reedited with the correct gRNA. Many examples of misediting/misguiding which are consistent with this hypothesis have been found (Fig. 9). However, another interpretation of unexpected patterns was also proposed (Decker and Sollner-Webb, 1990). It was suggested that editing is completely random and occurs between every nucleotide within an editing domain, and that when the correct sequence is formed it is frozen by the formation of base pairs with the gRNA. This issue will not be resolved until we have a complete knowledge of the total gRNA content of the mitochondrion and can compare all unexpected patterns to known gRNA sequences.

The enzyme cascade model is consistent with most observations, including the known 3’ to 5’ polarity of editing, but it does not satisfactorily explain the existence of the oligo-[U] tail on the gRNA. The role of the oligo-[U] tail was initially proposed to be a stabilization of the initial hybrid since the U’s would form base pairs with the G’s and A’s in the preedited region (Blum and Simpson, 1990). However, in 1991, we proposed that perhaps the oligo-[U] tail played a more active role and actually was the source of the U’s added during editing (Blum et al., 1991). A model was suggested in which the 3’ terminal OH of the gRNA attacked a phosphate within the mRNA at the site of the first

---

**TABLE 2**

Guide RNA Complexity in *L. tarentolae* UC and LEM125 Strains

<table>
<thead>
<tr>
<th>NUMBER OF gRNAs ENCODED BY CRYPTOGENE</th>
<th>maxicircle DNA</th>
<th>minicircle DNA</th>
<th>TOTAL (EXPECTED)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CRYPTOGENE</strong></td>
<td>UC + LEM125</td>
<td>UC + LEM125</td>
<td></td>
</tr>
<tr>
<td>COII</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>COIII</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ND7</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Cyb</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>MURF2</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>MURF4 (A6)</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>RPS12 (G6)</td>
<td>1</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td><strong>CRYPTOGENE</strong></td>
<td>UC + LEM125</td>
<td>LEM125</td>
<td></td>
</tr>
<tr>
<td>ND8 (G1)</td>
<td>1</td>
<td>9</td>
<td>9 (14)</td>
</tr>
<tr>
<td>ND9 (G2)</td>
<td>1</td>
<td>8</td>
<td>9 (17)</td>
</tr>
<tr>
<td>G3</td>
<td>2</td>
<td>1</td>
<td>3 (6)</td>
</tr>
<tr>
<td>G4</td>
<td>1</td>
<td>9</td>
<td>10 (15)</td>
</tr>
<tr>
<td>ND3 (G5)</td>
<td>1</td>
<td>5</td>
<td>6 (9)</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>13</td>
<td>47</td>
<td>60 (83)</td>
</tr>
</tbody>
</table>

*a* gND8-XII, a putative maxicircle-encoded gRNA with several mismatches.

*b* gG3-II, a putative maxicircle-encoded gRNA which was not detectable by Northern and primer extension analysis.

*c* gM150, a putative gRNA found in a gRNA–mRNA misguided chimera.
mismatch between the gRNA and mRNA, resulting in a transesterification and the exchange of the OH for the phosphate (see Fig. 7). The chemistry of this reaction is similar to that which occurs in self-splicing of RNA molecules in other cells. A similar transesterification model was independently proposed by Cech (1991). A prediction of this model is the existence of intermediate chimeric molecules which consist of gRNAs covalently linked to mRNAs at editing sites by the 3′ oligo-[U] tail. We searched for and found these chimeric molecules for three genes (Blum et al., 1991). This was gratifying but did not prove the transesterification model since chimeric molecules could possibly be
the stereochemistry of a gRNA-independent extract (Seiwert et al., 1994). The origin of editing was pushed back to the ancestor of the entire kinetoplastid lineage with the discovery of 5′-edited crypto-G1 mRNAs during a selective advantage for retention of this genetic system belongs to the Bodonina that represents another suborder within the kinetoplastids (Maslov and Simpson, 1994; Lukes et al., 1994). The occurrence of editing in the mitochondrion of Euglena is still an open question, although we showed (Yasuhira and Simpson, 1996) that the only mitochondrial gene cloned to date, COI, is unedited, and that gRNA-like molecules could not be detected in this organism by 5′ end capping with [α-32P]GTP and vaccinia virus guanylyl transferase. Since a phylogenetic analysis of the COI and the nuclear-encoded Hsp60 mitochondrial genes indicated a monophyletic origin of the Euglena and trypanosome mitochondria (Yasuhira and Simpson, 1996), a definitive demonstration of an absence of editing in Euglena mitochondria would indicate that editing is a derived character within the kinetoplastid lineage. A solution to this question awaits the analysis of additional mitochondrial genes from this organism.

Within the kinetoplastid lineages, the limitation of the editing of all known partially edited genes to the 5′ portions of editing domains and the presence of pan-edited homologs in related species suggest that 5′-edited cryptogenes, such as the COIII and ND7 genes in Leishmania and Crithidia, resulted from retroposition of cDNAs of partially edited mRNAs replacing the original pan-edited cryptogenes in the maxicircle genome (Simpson and Maslov, 1994) (Fig. 10). This implies that this type of RNA editing is a labile genetic trait which is easily lost in evolution. The observed loss of multiple minicircle sequence classes encoding gRNAs for the editing of the G1–G5 mRNAs during the prolonged culture history of the UC lab strain of L. tarentolae is entirely consistent with this hypothesis (Thiemann et al., 1994). This suggests that there must have been a selective advantage for retention of this genetic system.
RNA Editing Is Used to Describe a Diverse Set of Phenomena in Different Organisms in which RNA Molecules Are Modified in Sequence after Transcription

An even more complex type of insertional editing has been reported for mitochondrial transcripts from the acellular slime mold, Physarum (Gott et al., 1993; Mahendran et al., 1994). In this organism, all transcripts of the mitochondrial genome, including rRNAs, tRNAs, and mRNAs, are modified mainly by the insertion of C’s but also by the insertion of U, G, and A nucleotides, dinucleotide insertions, and even C to U substitutions. Little is known about the mechanism or mechanisms of these editing events.

Another type of editing was discovered in a mammalian nuclear gene for apolipoprotein B and also in plant mitochondria and chloroplasts, in which C’s are changed to U’s during the evolution of these cells. The selective pressure may be related to the fact that editing, at least in the complex life cycle of T. brucei, is regulated (Stuart, 1993). Editing is utilized by these cells as a translational control mechanism to control the biosynthesis of the mitochondrion. However, this does not explain the retention of editing in the monogenetic insect trypanosomes such as Crithidia, unless there are occult stages that undergo mitochondrial repression and derepression as in the digenetic African trypanosomes. This remains an open question.

If the mechanism of RNA editing in trypanosomes is determined to be unique to the parasite and not present in the human cell, then this pathway would be an excellent target for chemotherapy. Drugs which could selectively affect editing enzymes could theoretically kill the parasites without affecting the human host. It is hoped that the practical spin-off of this research on this bizarre genetic phenomenon may someday prove useful in treatment of the many trypanosome-caused human and animal diseases in Third World countries.
at precise sites (Hiesel et al., 1989; Covello and Gray, 1989; Gray and Covello, 1993). This was termed substitutional editing to distinguish it from the insertion/deletion type of editing in trypanosomes and Physarum. The C → U changes appear to involve a deamination of existing C nucleotides, but the way in which this is limited to multiple specific sites in the genome is not well understood, except in the case of the single apoB editing event in which specific proteins recognize short DNA sequences adjacent to the editing site (Backus and Smith, 1992, 1994).

Substitutional editing of tRNAs was also found to occur in mitochondria of the lower euukaryote, Acanthamoeba castellani, and in mitochondria of land snails: Single mismatched nucleotides within the acceptor stem were substituted by nucleotides which could form base pairs (Lonergan and Gray, 1993). In mitochondria of marsupials and rats, a unique type of substitutional editing of tRNA was reported in which a single C to U change occurred within or near the anticodon sequence (Morl et al., 1995).

Another class of substitutional editing was found to occur in several glutamate receptor mRNAs in humans and involves the deamination of A residues into I residues, which are treated by the translational machinery as G residues. The A to I type of editing of the glutamate receptor mRNA in mammals was discovered in an unusual way. Initially, an enzyme activity was described which could unwind double-stranded RNAs by changing A into I residues (Bass and Weintraub, 1988). A double-strand RNA adenosine deaminase (dsRAD or DRADA) which catalyzes the deamination of A into I has been purified and the gene cloned (Kim et al., 1994; Wang et al., 1995). A biological function was finally found for this activity when it was reported that glutamate receptor mRNAs were modified by A to G changes at specific sites, and that the modifications required the formation of a RNA duplex by the foldback of a complementary downstream intron sequence (Higuchi et al., 1993; Mans et al., 1996). Since the translational machinery treats Is as Gs, dsRAD was immediately suspected to be the culprit. It now appears that another related adenosine deaminase, RED1, is responsible for this specific example of editing (Melcher et al., 1996). However, dsRAD has been implicated in the A to G editing event in hepatitis B delta viral antigenic RNA. The A to I conversion editing is likely to be quite widespread due to the ubiquity of dsRAD-like activities in higher organisms.

Another modification of RNA, which has been termed editing, involves the addition of Gs in mRNAs of certain RNA viruses (Vidal et al., 1990; Pelet et al., 1991; Curran et al., 1991). In several paramyxoviruses a unique P gene gives rise to two mRNAs. One is a faithful copy of the DNA and the other contains one or two extra Gs inserted within a run of five or six Gs. The resulting frameshifts allow ribosomal access to a second downstream reading frame, resulting in an alternate P protein with a different C-terminal sequence. The addition of Gs in the viral mRNAs is probably due to “stuttering” of the RNA during transcription.

Conclusions

The modification of RNA sequences after transcription is a widespread phenomenon among euukaryotic cells and involves several different types of mechanisms. The trypanosome mitochondrial U-insertion/deletion type of editing is apparently unique to the kinetoplastid protozoa. The Physarum mitochondrial C-insertion editing has some similarities to the trypanosome editing, but more information is required to make any conclusions. The plant mitochondrial and apo B C to U substitutional editing and the glutamate receptor A to I editing both involve deaminations, but the structural requirements, catalytic mechanisms, and RNA binding sites are different for the two reactions. One common feature between the A to I editing and the U-insertion/deletion editing is a requirement for double-stranded RNA as a recognition element.

The biological importance of the various types of RNA editing is emphasized by the fact that these modifications are frequently regulated and have significant biological phenotypic consequences. It is clear that additional examples of RNA editing will be uncovered in the future, especially as entire genomic sequences of multiple organisms become available.

References Cited

LARRY SIMPSON

four nucleotides that are not encoded in the DNA. *Cell* **46**, 819–826.


General References


