Letter to the Editor

Introns: Mighty Elements from the RNA World

Alexei Fedorov, Larisa Fedorova

Department of Medicine, Medical College of Ohio, Toledo, OH 43614, USA

Abstract. The discovery of RNA-based enzymatic activity by Thomas Cech's and Sidney Altman's laboratories was a momentous event that led Walter Gilbert to the concept of an ''RNA world''—a primitive ancient stage of life that existed before the appearance of DNA and protein molecules. A year later, Gilbert formulated ''the exon theory of genes,'' which hypothesized that introns are very ancient genetic elements present at the earliest stages of life in the RNA world. This theory has been fiercely debated and still has vigorous supporters and opponents. In this communication, we explore peculiarities in the RNA-protein world and their effect on intron–exon structures. We demonstrate that these peculiarities, which exist in the absence of DNA, could shed light on introns' original functions as well as the important role they might have played in the origin of life. For ancient DNA-lacking cells, a crucial problem existed in distinguishing two distinct subsets of RNAs: those messenger molecules coding for proteins and those heritable genetic molecules complementary to messenger RNAs that propagate the genetic information through generations. We propose that ancient introns could act as markers of RNA subsets, directing them to different functions.

Key words: Intron — Exon — RNA world — Function — Evolution — Origin — Genome

Correspondence to: Alexei Fedorov

Introduction

Introns are widespread among all eukaryotes (Lynch and Richardson 2002; Fedorova and Fedorov 2003), while only a limited number of group I and group II introns were found in prokaryotes (Shub 1991; Martinez-Abarca and Toro 2000). Whether or not introns had ever been abundant in bacteria was the most debatable question among evolutionists studying exon–intron gene structures. This question gave rise to two opposing theories known as introns-early and introns-late. The introns-late theory holds that introns appeared relatively late in eukaryote evolution as selfish transposable elements and have never been abundant in bacteria (Cavalier-Smith 1985; Palmer and Logsdon 1991; Logsdon et al. 1998). Thus, introns had no important function at the time of their emergence. By contrast, the introns-early theory supposes that introns already existed at the origin of life during the RNA world and played a very important role by facilitating exon shuffling (Gilbert 1987; Long et al. 1995, 2003). Exon shuffling, which combines exons from different genes, eventually led to the assembly of long modern genes from a limited number of ''pieces'' constituting ancient minigenes. However, opponents of introns-early argue that successful exon shuffling events would have been exceedingly rare, happening once in, perhaps, several million years. Thus, exon shuffling could be considered insufficient in explaining the expansion of introns. Furthermore, Lynch and Richardson (2002) argue in a recent review that introns create several problems for their host cells. If there is no powerful force driving the proliferation of introns and they are harmful to their host genes, ''*How, then, did the*

eukaryotic genome arrive at the point at which 'genes in pieces' has become the norm?'' (Lynch and Richardson 2002). Here we propose an answer to this fundamental question. We point out that introns could act as extremely valuable signals for marking specific subsets of nucleic acid molecules and governing their fate in the cells.

We investigate possible roles of introns during the earliest stages of evolution in the RNA world (Gilbert 1986). One of the pioneers of RNA world, Leslie Orgel (2003), recently wrote, "*It is now generally accepted that our familiar biological world was preceded by an RNA world.*" However, we do not know many details of this world and when it was replaced by modern DNA-containing organisms. Initially, the RNA world was hypothesized as being limited to the earliest stages of life, existing before the divergence of eukaryotes and prokaryotes. However, several bioinformatic investigations suggest that RNA-based primordial cells were involved in the origin of eukaryotes (Sogin 1991; Doolittle 1995; Hartman and Fedorov 2002). Some authors argue for the antiquity of eukaryotes, placing them at the root of the universal tree of life (Gribaldo and Phillippe 2002). Thus, a discussion of the presence of introns in the RNA world neither is beholden to one particular view of ancient evolution nor is it necessarily in contradiction to the introns-late theory. We leave aside the difficult question pertaining to the origin of eukaryotes and focus instead on the potential usefulness of introns in the RNA world.

Mighty Introns Hypothesis

It is widely accepted that, at the stage of the RNA-protein world, there existed multiple minigenes—short RNA molecules (similar to present mRNAs) which encoded individual proteins. At first glance, one would expect that all introns would have been immediately excised from these minigenes via splicing and should, therefore, have vanished from the face of the earth. However, the complementary RNA copy of an intron does not have splicing properties and is hence enzymatically nonactive. Let us call this form as "nonactive complementary introns." To multiply and propagate through generations, minigenes must have their complementary templates, which we call "minus-strand replicas." Thus introns in the RNA-protein world would be continuously present only in the form of nonactive complementary strands within minus-strand replicas. Transcription of minus-strand replicas should produce minigenes with active introns, which would be immediately spliced out as discussed earlier (Fig. 1).

Contemporary genomic analyses show that numerous introns already existed at the earliest stages of eukaryote evolution (Rogozin et al. 2003; Fedorov et al. 2002). In addition, in every species of eukaryotes there is a very close association of splicing with other chemical modifications of mRNA: capping at the 5′ end and poly(A) adenylation at the 3′ end (Proudfoot et al. 2002). Poly(A) adenylation is specific to all eukaryotes, while 5′-capping exists also in the bacterial world. It looks reasonable that these two types of mRNA modifications are at least as ancient as splicing. We make a sole speculation for our hypothesis that association of splicing and chemical modifications of mRNA ends already existed in the RNA-protein world. In our scenario, the presence of introns in the RNA-protein world would produce significant nonsymmetry between the strands: Minus-strand replicas would have nonactive complementary introns and mini-genes would be intron-less and modified from both ends due to splicing. We suggest that this nonsymmetry was crucial for the RNA-protein world. In particular, this nonsymmetry ensures that (i) minus-strand replicas would not be translated because they do not have chemical modifications on both ends and (ii) minigenes would not be transcribed because they do have modifications. Moreover, the removal of introns from minigenes also prevents complete renaturation of minigenes with their replicas and allows production of multiple minigene copies from a single replica molecule (Fig. 1).

Splicing is another area that merits further discussion. The splicing process is active only under specific conditions—presence of splicing proteins or optimal concentration of ions. Hence, it is possible to halt the splicing temporarily by inactivating the spliceosomal protein(s) or by changing the ion concentration. This splicing inactivation results in a production of intron-containing minigenes without chemically modified termini. Absence of the-5′-cap and poly(A) tail of these molecules would be a signal for their usage as templates for replication. The transcription of these intron-containing minigenes would multiply the number of minus-strand replicas, a step required for cell division. After this replication cycle, cell conditions would be restored to favorable of splicing, and intron-containing minigenes would undergo splicing, producing minigenes suitable for protein translation. Therefore, the regulation of splicing could be the major regulator of the cell cycle in the RNA-protein world. This hypothesis has some corroboration in present-day retroviruses with RNA genomes. All members of retroveridae viruses do have introns (Coffin 1996) and efficient replication of their genomes requires a precise balance between the unspliced and the spliced RNA (McNally and McNally 1998).

From the theoretical viewpoint, life is a specific set of chemical processes generating and maintaining
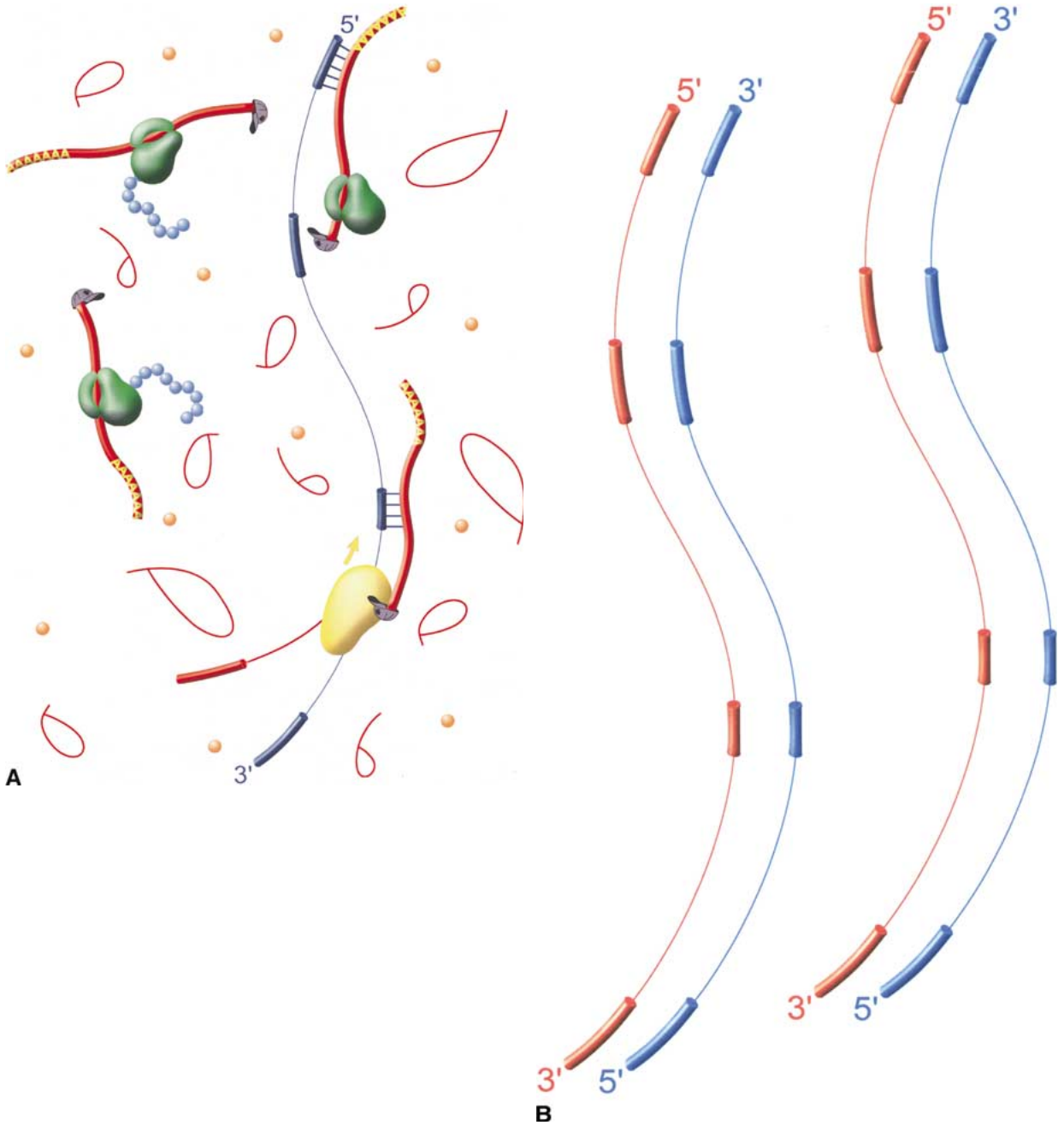
Fig. 1. Splicing in the RNA-protein world. Minigenes are shown in red, and minus strand replicas in blue. Introns and nonactive complementary introns are shown as thin lines; exons and their complementary strands, as cylinders. Cellular cofactors required for splicing are presented as small orange circles. **A** Conditions favorable to splicing. Green particles, ribosomes; yellow ellipse, RNA–RNA polymerase; chains of blue beads, proteins; red loops, excised introns. **B** Conditions unfavorable to splicing (RNA-genome replication stage).

chemical disequilibrium. In the RNA-protein world, it is of ultimate importance to have a mechanism for distinction of protein coding RNA strands (minigenes) and their complementary templates (minus-stand replicas). The presence of introns resolves this problem, creating disequilibrium between the different strands of RNA. Analogous distinction of two kinds of DNA molecules exists in ciliates, single-cell eukaryotes. The macronucleus of ciliates contains short DNA sequences of transcriptionally active genes, while their micronucleus contains much larger and transcriptionally inactive DNA molecules used for inheritance of genetic material. Interestingly, many micronuclear genes are interrupted by intervening noncoding sequences, also known as internal eliminated segments (IESs) (Prescott 1999). In analogy to introns, IESs interrupt gene coding sequences into several parts, yet they are excised from the micronucleus genes by a completely different mechanism on a DNA level.

In summary, we propose that introns and other noncoding segments of nucleic acids could be mighty functional elements governing the fate of RNA and DNA molecules in the RNA-protein world as well as more advanced species.

References

Coffin JM (1996) Retroviridae: The viruses and their replication. In: Fields BN, Knipe DM, Howley PM (eds) Fields virology. Raven Press, New York, pp 1767–1847

Doolittle RF (1995) The origin and evolution of eukaryotic proteins. Philos Trans R Soc Lond B Biol Sci 349:235–240

Fedorov A, Merican AF, Gilbert W (2002) Large-scale comparison of intron positions among animal, plant, and fungal genes. Proc Natl Acad Sci USA 99:16128–16133

Fedorova L, Fedorov A (2003) Introns in gene evolution. Genetica 118:123–131

Gilbert W (1986) Origin of life, the RNA world. Nature 319:618

Gilbert W (1987) The exon theory of genes. Cold Spring Harbor Symp Quant Biol 52:901–905

Gribaldo S, Phillippe H (2002) Ancient phylogenetic relationships. Theor Popul Biol 61:391–408

Hartman H, Fedorov A (2002) The origin of the eukaryotic cell: A genomic investigation. Proc Natl Acad Sci USA 99:1420–1425

Logsdon JM Jr, Stoltzfus A, Doolittle WF (1998) Molecular evolution: Recent cases of spliceosomal intron gain? Curr Biol 8:R560–R563

Long M, deSouza SJ, Gilbert W (1995) Evolution of exon-intron gene structure of eukaryotic genes. Curr Opin Genet Dev 5:774–778

Long M, Deutsch M, Wang W, Betran E, Brunet FG, Zhang J (2003) Origin of new genes: evidence from experimental and computational analyses. Genetica 118:171–182

Lynch M, Richardson AO (2002) The evolution of spliceosomal introns. Curr Opin Genet Dev 12:701–710

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

McNally LM, McNally MT (1998) An RNA splicing enhancer-like sequence is a component of a splicing inhibitor element from Rous Sarcoma virus. Mol Cell Biol 18:3103–3111

Orgel LE (2003) Some consequences of the RNA world hypothesis. Orig Life Evol Biosph 33:211–218

Palmer JD, Logsdon JM (1991) The recent origin of introns. Curr Opin Genet Dev 1:470–477

Prescott DM (1999) The evolutionary scrambling and developmental unscrambling of germline genes in hypotrichous ciliates. Nucleic Acids Res 27:1243–1250

Proudfoot NJ, Furger A, Dye MJ (2002) Integrating mRNA processing with transcription. Cell 108:501–512

Rogozin IB, Wolf YI, Sorokin AV, Mirkin BG, Koonin EV (2003) Remarkable interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in eukaryotic evolution. Curr Biol 13:1512–1517

Shub DA (1991) The antiquity of group I introns. Curr Opin Genet Dev 1:478–484

Sogin ML (1991) Early evolution and the origin of eukaryotes. Curr Opin Genet Dev 1:457–463