J Mol Evol (2012) 74:226–231
DOI 10.1007/s00239-012-9494-3

Selective forces for the origin of spliceosomes

Matej Vesteg · Zuzana Šándorová · Juraj Krajčovič

Received: 3 October 2011 / Accepted: 24 February 2012 / Published online: 11 March 2012
© Springer Science+Business Media, LLC 2012

Abstract It has been proposed that eukaryotic spliceosomes evolved from bacterial group II introns via constructive neutral changes. However, a more likely interpretation is that spliceosomes and group II introns share a common undefined RNA ancestor—a proto-spliceosome. Although, the constructive neutral evolution may have probably played some roles in the development of complexity including the evolution of modern spliceosomes, in fact, the origin, losses and the retention of spliceosomes can be explained straight-forwardly mainly by positive and negative selection: (1) proto-spliceosomes evolved in the RNA world as a mechanism to excise functional RNAs from an RNA genome and to join non-coding information (ancestral to exons) possibly designed to be degraded. (2) The complexity of proto-spliceosomes increased with the invention of protein synthesis in the RNP world and they were adopted for (a) the addition of translation signal to RNAs via *trans*-splicing, and for (b) the exon-shuffling such as to join together exons coding separate protein domains, to translate them as a single unit and thus to facilitate the molecular interaction of protein domains needed to be assembled to functional catalytic complexes. (3) Finally, the spliceosomes were adopted for *cis*-splicing of (mainly) non-coding information (contemporary introns) to yield translatable mRNAs. (4) Spliceosome-negative organisms (i.e., prokaryotes) have been selected in the DNA–protein world to save a lot of energy. (5) Spliceosome-positive organisms (i.e., eukaryotes) have been selected, because they have been completely spliceosome-dependent.

Keywords Constructive neutral evolution · Ribosome · RNA world · Selection · *Trans*-splicing

Abbreviations
CNE Constructive neutral evolution
LECA Last eukaryotic common ancestor
LUCA Last universal common ancestor
RNP Ribonucleoprotein

Some reports have appeared in the recent literature advocating constructive neutral evolution (CNE) to explain the evolution of apparently complex cellular systems (Doolittle et al. 2011; Gray et al. 2010; Lukeš et al. 2011) based on the previous hypotheses of Covello and Gray (1993) and Stoltzfus (1999). Constructive neutral model suggests that a neutral mutation can allow the retention of a subsequent evolutionary change which would be otherwise lethal or that even slightly deleterious mutation can be compensated by another subsequent evolutionary change (Gray et al. 2010; Lukeš et al. 2011; Stoltzfus 1999). The repeated cycles of such stochastic evolutionary changes can lead to the addition of more and more macromolecules or their fragments to the system which becomes completely dependent on these components and thus the complexity of a cellular process can increase (Gray et al. 2010; Lukeš et al. 2011; Stoltzfus 1999).

M. Vesteg (✉) · Z. Šándorová · J. Krajčovič
Department of Genetics, Faculty of Natural Sciences, Comenius University, Mlynská dolina B-1, 842 15 Bratislava, Slovakia
e-mail: vesteg@fns.uniba.sk; krajcovic@fns.uniba.sk

M. Vesteg · J. Krajčovič
Institute of Cell Biology and Biotechnology, Faculty of Natural Sciences, Comenius University, Bratislava, Slovakia

Doolittle et al. (2011), Gray et al. (2010), Lukeš et al. (2011) and Stoltzfus (1999) suggest that the origin and evolution of huge molecular complexes such as ribosomes, spliceosomes, and editosomes can be mainly explained by CNE. Nevertheless, although CNE may have broad evolutionary implications, it is questionable if CNE can be equally applied for explaining the origins of these three complex structures in which RNAs exhibit their crucial function. While the origin of ribosomes dates certainly back to the last universal common ancestor (LUCA) (Woese 1998, 2000, 2001, 2002) and the spliceosomes were certainly present in the last eukaryotic common ancestor (LECA) (Collins and Penny 2005), the editosomes were invented only recently, probably in the ancestor of kinetoplastid lineage (Flegontov et al. 2011). Moreover, while catalytic ribosomal rRNAs are probably the most ancient ribozymes which still exist—they are clearly relics of the RNA world (Cech 2011), and the key catalytic function of spliceosomes is achieved by five cross-interacting snRNAs (Valadkhan and Jaladat 2010), gRNAs involved in RNA editing in kinetoplastid mitochondria just hybridize to pre-mRNAs (they serve only as a template for the correction of dismissed information) and they have no catalytic function—gRNAs are not ribozymes (Panigrahi et al. 2006). Finally, the spliceosomal RNP complex is almost as large as eukaryotic large ribosomal subunit (up to 60S) (Valadkhan and Jaladat 2010), while the editosome is a unique 20S complex composed solely of proteins (Panigrahi et al. 2006). In addition, the genomes of intracellular endosymbionts (such as mitochondria) are affected by higher mutation frequency, load, and drift and thus ratchet-like processes in comparison to nuclear genomes (Andersson et al. 2002; Wernegreen 2005; Woolfit and Bromham 2003). All common characteristics of spliceosomes and ribosomes and their differences from characteristics of editosomes suggest that the mechanisms for the origin and evolution of editosomes were partly different from those shaping the origin and evolution of ribosomes and spliceosomes. On the other hand, it is sure that ribosomes are obligatory for all organisms, spliceosomes are obligatory for all eukaryotes and editosomes are probably obligatory for the most of kinetoplastids. Hence, there is currently strong positive selective pressure to retain (or strong negative selective pressure to loose) these huge molecular complexes, because the survival of their carriers completely depends on them. It is also likely that a strong positive (and/or negative) selection largely shaped the origin and evolution of these molecular complexes in any transient stage, even if the function of a transient stage was different from their contemporary function.

Doolittle et al. (2011), Gray et al. (2010), Lukeš et al. (2011) and Stoltzfus (1999) suggest that spliceosomal introns and spliceosomes evolved by CNE from bacterial self-splicing group II introns which are often assumed to have been introduced to a prokaryotic ancestor of eukaryotes by the α-proteobacterial ancestors of mitochondria. This scenario is based on the proposal that prokaryotes (i.e., archaea) were ancestral to eukaryotes. However, no direct evidence supporting such notion exists. Instead, the current knowledge of biochemical properties of membranes, molecular phylogenies, distribution of genes among organisms, cell biology, and population genetics is rather consistent with the view that all organisms share an undefined common ancestor (LUCA) and that the root of the universal phylogenetic tree is between the Bacterial domain and the rest of organisms—the domains Archaea and Eukarya which share a more recent common ancestor (for review see Vesteg and Krajčovič 2011). With this respect, it is likely that bacterial group II introns share a common undefined ancestor with spliceosomal introns a spliceosomes—a proto-spliceosome.

Proposing that group II introns were ancestral to spliceosomal ones is not neutral (nor slightly costly) but seriously a costly scenario. It has been known for more than two decades that spliceosome-mediated splicing is slower than splicing of group II introns (Baurén and Wieslander 1994; Beyer and Osheim 1988; Lang and Spritz 1987; Schmidt et al. 1996). Although group II introns are among the largest and slowest ribozymes whose folding to the native structure takes approximately 30 s in vitro, they are spliced much faster in vivo in the presence of protein co-factors (Fedorova et al. 2010; Pyle et al. 2007; Roitzsch and Pyle 2009). In contrast, in vitro and in vivo experiments suggest that spliceosomal introns are excised in approximately 1–20 min (Audibert et al. 2002; Kessler et al. 1993; Singh and Padgett 2009). Lukeš et al. (2011) argue that five spliceosomal RNAs are unlikely to get together to make a group II intron. However, the spliceosomal RNAs are the components of the same RNP complex processing genetic information. Hence, the possibility that these RNAs would be reversely transcribed and inserted into the host DNA genome tightly linked together does not seem to be so unlikely. Their further reduction to group II introns does not appear to be an unlikely proposal as well, because group II introns are just faster. As we have recently argued (Vesteg and Krajčovič 2011), the evolution of slower and more complex RNA-based catalytic system from faster and simpler one could be hardly considered as being selectively neutral (nor nearly neutral) in every transient stage. In addition, much more efficient purely protein-based catalytic system would be predicted to evolve from simple RNA-based system rather than a less efficient RNA-based system requiring interaction and catalysis of various RNA molecules. For example, the protein complexity of mitochondrial ribosomes has increased in comparison to the ribosomes of their bacterial ancestors, while the rRNA complexity has decreased (Lukeš et al. 2011).

Therefore, the RNA complexity of the spliceosome would be also expected to decrease in evolution, while the spliceosomes-late scenario proposes exactly the opposite. Moreover, the expansion of group II introns from the α-proteobacterial ancestors of mitochondria would be either impossible in an asexual hypothetical prokaryotic ancestor of eukaryotes or it would be dramatically fitness-reducing, because it would slow down both DNA replication and gene expression of a hypothetical prokaryotic host (for review see Vesteg and Krajčovič 2011). Since the spliceosomes-late scenario avoids any adaptive reason to explain the origin of the spliceosome, it can be considered as a virtual creation of less efficient and more complex RNA-based system from more efficient and simpler RNA-based system in the DNA–protein world in which enzymes are approximately million times more efficient than ribozymes (Jeffares et al. 1998; for review see Penny et al. 2009). This model is definitely not neutral and it would need to suggest clear selective advantages compensating for the serious costs to become plausible.

In our opinion, a proto-spliceosome could have likely been more complex than group II introns and less complex than contemporary spliceosome, while it could have been less efficient and slower than both these RNA-based machines. The selection for higher efficiency of gene expression, and higher speed and fidelity of splicing could have been the main evolutionary forces shaping the origin of both spliceosomes and group II introns. However, the diversification of these two RNA-based structures from their common ancestor could have probably occurred in different populations of cells living in different conditions under different selective pressures.

Although slower spliceosomes need not to be necessarily disadvantageous in comparison to faster group II introns, if gene regulation and expression control become more critical than speed during the advent of complexity, the key problem with the proposal for the origin of spliceosomal introns from group II introns is that this scenario largely ignores the RNA continuity theory—the proposal that the ancestry of most (if not all) catalytic RNAs dates back to the RNA world. The ancestors of these catalytic RNAs could have been initially likely selected for different function from that of their descendants. Nevertheless, the ancient ribozymes were likely pre-adapted and later selected (adopted) for similar function like they achieve presently (Penny et al. 2009; Poole et al. 1999). Although the regulatory or other non-catalytic short RNAs are expected to arise also after the invention of translation and genetic code, the ribozymes (whose catalysis is 10^3–10^9 times less efficient than that of enzymes) are not expected to evolve and to substitute for the catalysis achieved already by the enzymes (Penny et al. 2009; Poole et al. 1999). Hence, the origin of spliceosomal snRNAs most likely dates back to the RNA world as well as the origin of rRNAs (Cech 2011).

Based on the universal presence of ribosomes, Poole et al. (1998, 1999) suggested that proto-ribosomes of the first riboorganisms might have been ancestrally involved in RNA replication which might have been mediated by tRNAs—the donors of complementary trinucleotides. The selective advantage of such a replication mechanism could have been the proofreading (or elimination of unnecessary and foreign RNA)—only those RNAs recognized by tRNAs were replicated by a step-wise addition of three nucleotides at once (Poole et al. 1999). The amino acids might have initially served only for stabilization of tRNA-SSU rRNA interaction and the LSU rRNA might have been only later integrated into the complex (Poole et al. 1999). The proto-ribosomes were thereafter, probably after the invention of the first peptides with the RNA-dependent RNA-polymerasing activity, fully adopted for translation. The evolution of such a proto-ribosome could have been mediated by the series of Darwin–Eigen hypercycles (Poole et al. 1999). A step-wise addition of proteins to the RNA backbone of a proto-ribosome could have led to the origin of a primitive ribosome specialized for translation. The primitive translation system of LUCA is proposed to have been ancestrally much more error-prone and a step-wise selection for higher efficiency and fidelity of translation is expected to have formed the structure of ribosomes (Woese 1998, 2000, 2001, 2002).

Similarly like proto-ribosome, the proto-spliceosome of the first riboorganisms might have ancestrally achieved completely different function from that of contemporary spliceosomes. The introns-first theory suggests that the genome of the first RNA-based organisms was composed of RNA modules encoding functional ribozymes ancestral to contemporary introns and a small portion of waste information ancestral to contemporary exons (Penny et al. 2009; Poole et al. 1999). The proto-spliceosomal RNAs could have been selected for their capability of the excision of functional ribozymes and joining together unnecessary information (the ancestor of mRNAs), eventually recognized, and degraded by other ribozymes. Following this scenario, the proto-spliceosomal RNAs could be viewed as selfish elements, since they excised themselves from host RNAs to yield their functional catalytic forms, but they were also altruistic and advantageous for their carriers, because they could have promoted maturation of other host ribozymes. It is also likely that the invention of primitive proteins contributed to the complexity of the proto-spliceosomes and that translation and splicing somehow co-evolved in a step-wise manner mediated by Darwin–Eigen hypercycles.

Concomitant with the evolution of genetic code and translation, the proto-spliceosomes could have been

adopted to reshuffle small-peptide encoding regions and thus to increase the variability of arising proteins—the hypothesis that splicing allowed recombination of RNA modules encoding protein domains (exon-shuffling) and thus it was adaptive (Doolittle 1978; Gilbert 1978; Kurland et al. 2007). If this scenario is considered purely as occurring exclusively at the DNA level via recombination of introns, it can be criticized as a creationistic scenario including a foresight (Blake 1978). However, the exon-shuffling at the RNA level (either via recombination or via *trans*-splicing) could have been rather an obligatory part of gene expression in the RNP-based LUCA. The spliceosome-dependent *trans*-splicing could have helped to bring together dispersed information and to increase the efficiency of gene expression. The exons could have ancestrally encoded independent protein domains which could have functioned in domain complexes—the hypothetical ancestors of modern proteins composed of various protein domains. Bringing the exons together by *trans*-splicing, translating them as a single unit, and producing a single protein containing many coupled domains, could have increased the efficiency of both gene expression and protein domain catalysis. The translation of many domains at once would at least minimize the time to assemble separate domains into catalytic complexes. Following this scenario, the increased variability of proteins amenable to selection would be only a side effect of the RNA exon-shuffling or spliceosome-dependent *trans*-splicing process which was ancestrally likely selected as an innovative and more parsimonious way of gene expression saving the time and energy in the RNP world.

It is also possible that the origin of spliced-leader (SL-)RNA-mediated spliceosome-dependent *trans*-splicing occurring in a variety of unrelated eukaryotes (for review see Hastings 2005) may trace back to the RNP-based LUCA. This process is obligatory for these organisms, because mRNAs obtain cap from 5'-end of SL-RNA via this process and the cap is necessary for translation. This process does not allow effective regulation of gene expression at the level of transcription initiation, because SL-RNAs and pre-mRNAs are transcribed from different genomic regions (Martínez-Calvillo et al. 2010; Vesteg et al. 2009). Therefore, the process is energetically costly in comparison to the effective regulatory condition seen in prokaryotes and many model eukaryotes. It has been proposed that SL-RNA-mediated *trans*-splicing evolved many times independently in eukaryotes (Derelle et al. 2010; Douris et al. 2010; Lukeš et al. 2009). However, assuming a secondary de novo appearance of such a costly process in the DNA–protein world even once appears to be highly improbable from the selectionistic point of view. The recent bioinformatic comparison of SL-RNAs from various unrelated eukaryotes revealed that their secondary structures are strikingly similar suggesting that all SL-RNAs share a common ancestor which was present in LECA (Marz et al. 2010) and likely also in LUCA.

The role for selfish replicative elements (i.e., viruses) in the origin and evolution of spliceosomes as well as ribosomes can also not be ruled out. Such viruses were most likely abundant in the RNA and the RNP world (Gilbert and de Souza 1999) and they were probably attacking the first riboorganisms all the time. The theoretical simulations suggest that altruistic features can easily appear in hypercyclic interactions of self-replicators and the hypercycle can easily become resistant against other parasites (Boerlijst and Hogeweg 1991). The experimental evidence also exists that a conflict mediation system (e.g., two competitive viruses invading the same host) can result in the evolution of a highly cooperative system in quite a short time (Sachs and Bull 2005).

DNA could have been invented by viruses, because it was more stable molecule than RNA and it would not be recognized by “immunity system” (RNA inactivation mechanisms) of an RNP-based host (Forterre 2002, 2005, 2006). Owing to the enhanced stability and thus coding capacity of DNA, DNA was later adopted by the RNP-based host(s), although it is questionable if such “DNA-transition” happened once in LUCA (Vesteg and Krajčovič 2008; Woese 1998, 2000, 2002), twice—in the ancestor of Bacteria, and in the common ancestor of Archaea and Eukarya (Forterre 2005; Poole and Logan 2005), or three times—in the ancestors of each domain independently (Forterre 2005, 2006).

Various selective forces for the simplification of cell structure in the ancestors of prokaryotes and for the origin of parsimoniously organized prokaryotic genomes (genes without spliceosomal introns are organized in operons) have been suggested: thermoreduction (Forterre 1995), r-selection (Poole et al. 1998), efficient selection in large populations (Lynch 2007), chronic energy stress (Valentine 2007), or other reasons (for review see Penny et al. 2009). All these selective mechanisms can be also applied for explaining the reduction of proto-spliceosomes to group II introns and almost complete intron loss in prokaryotes. On the other hand, the persistence of spliceosomes in eukaryotes can be explained by complete dependency of eukaryotes and their ancestors (LECA) on this RNP complex. An intermediary evolutionary stage analogous to that leading to the origin prokaryotic operonal gene structure from eukaryotic-like one can be exemplified by trypanosomatids. In contrast to their relatives—free-living euglenids (Vesteg et al. 2010), trypanosomatid parasites are almost completely devoid of *cis*-spliceosomal introns. Nevertheless, trypanosomatid pre-mRNAs are polycistronic containing transcripts of genes separated by *trans*-introns and the individual mRNAs are resolved and capped by spliceosome-dependent

SL-RNA-mediated *trans*-splicing (for review see Martínez-Calvillo et al. 2010). In this case, the origin of prokaryotic gene expression would require nothing else than replacement (or reduction) of *trans*-introns by (to) something like Shine–Dalgarno sequence (ribosome binding site) and its recognition by ribosome accompanied by the loss of nuclear membrane, the loss of spliceosomes and SL-RNAs, and the reduction of ribosomes.

Taken altogether, although CNE could have somehow contributed to the origin and evolution of spliceosomes and ribosomes in any transient stage described here, there is no reason to suggest it as the dominant mechanism. The origin and evolution of these huge molecular ribomachines can be largely explained by applying selectionistic criteria.

Acknowledgments This work was supported by a grant VEGA 1/0416/09 from the Ministry of Education of the Slovak Republic to J. K., and is the result of the project implementation: "the improvement of centre of excellence for exploitation of informational biomacromolecules in disease prevention and improvement of quality of life," ITMS 26240120027, supported by the Research and Development Operational Programme funded by the ERDF.

References

Andersson SG, Alsmark C, Canbäck B, Davids W, Frank C, Karlberg O, Klasson L, Antoine-Legault B, Mira A, Tamas I (2002) Comparative genomics of microbial pathogens and symbionts. Bioinformatics 18:S17

Audibert A, Weil D, Dautry F (2002) In vivo kinetics of mRNA splicing and transport in mammalian cells. Mol Cell Biol 22:6706–6718

Baurén G, Wieslander L (1994) Splicing of balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. Cell 76:183–192

Beyer AL, Osheim YN (1988) Splice site selection, rate of splicing and alternative splicing on nascent transcripts. Genes Dev 2:754–765

Blake CC (1978) Do genes-in-pieces imply proteins-in-pieces? Nature 273:267

Boerlijst MC, Hogeweg P (1991) Spiral wave structures in prebiotic evolution—hypercycles stable against parasites. Phys D 48:17–28

Cech TR (2011) The RNA world in context. Cold Spring Harb Perspect Biol. doi:10.1101/cshperspect.a006742

Collins L, Penny D (2005) Complex spliceosomal organization ancestral to extant eukaryotes. Mol Biol Evol 22:1053–1066

Covello PS, Gray MW (1993) On the evolution of RNA editing. Trends Genet 9:265–268

Derelle R, Momose T, Manuel M, Da Silva C, Wincker P, Houliston E (2010) Convergent origins and rapid evolution of spliced leader *trans*-splicing in metazoa: insights from the ctenophora and hydrozoa. RNA 16:696–707

Doolittle WF (1978) Genes in pieces: were they ever together? Nature 272:581–582

Doolittle WF, Lukeš J, Archibald JM, Keeling PJ, Gray MW (2011) Comment on "Does constructive neutral evolution play an important role in the origin of cellular complexity?". BioEssays 33:427–429

Douris V, Telford MJ, Averof M (2010) Evidence for multiple independent origins of *trans*-splicing in metazoa. Mol Biol Evol 27:684–693

Fedorova O, Solem A, Pyle AM (2010) Protein-facilitated folding of group II intron ribozymes. J Mol Biol 397:799–813

Flegontov P, Gray MW, Burger G, Lukeš J (2011) Gene fragmentation: a key to mitochondrial genome evolution in euglenozoa? Curr Genet 57:225–232

Forterre P (1995) Thermoreduction, a hypothesis for the origin of prokaryotes. C R Acad Sci Paris Life Sci 318:414–422

Forterre P (2002) The origin of DNA genomes and DNA replication proteins. Curr Opin Microbiol 5:525–532

Forterre P (2005) Two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. Biochimie 87: 793–803

Forterre P (2006) Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domain. Proc Natl Acad Sci USA 103:3669–3674

Gilbert W (1978) Why genes in pieces? Nature 271:501

Gilbert W, de Souza SJ (1999) Introns and the RNA world. In: Gesteland RF, Cech TR, Atkins RF (eds) The RNA world, the nature of modern RNA suggests a prebiotic RNA World, vol. 37, 2nd edn. CSHL Press, New York, pp 221–231

Gray MW, Lukeš J, Archibald JM, Keeling PJ, Doolitle WF (2010) Cell biology. Irremediable complexity? Science 12:920–921

Hastings KEM (2005) SL *trans*-splicing: easy come or easy go? Trends Genet 21:240–247

Jeffares DC, Poole AM, Penny D (1998) Relics from RNA world. J Mol Evol 46:18–36

Kessler O, Jiang Y, Chasin LA (1993) Order of intron removal during splicing of endogenous adenine phosphoribosyltransferase and dihydrofolate reductase pre-mRNA. Mol Cell Biol 13:6211–6222

Kurland CG, Canbäck B, Berg OG (2007) The origins of modern proteomes. Biochimie 89:1454–1463

Lang KM, Spritz RA (1987) In vitro splicing pathways of pre-mRNAs containing multiple intervening sequences. Mol Cell Biol 7:3428–3437

Lukeš J, Leander BS, Keeling PJ (2009) Cascades of convergent evolution: the corresponding evolutionary histories of euglenozoans and dinoflagellates. Proc Natl Acad Sci USA 106: 9963–9970

Lukeš J, Archibald JM, Keeling PJ, Doolittle WF, Gray MW (2011) How a neutral evolutionary ratchet can build cellular complexity. IUBMB Life 63:528–537

Lynch M (2007) The origins of genome architecture. Sinauer, Sunderland

Martínez-Calvillo S, Vizuet-de-Rueda JC, Florencio-Martínez LE, Manning-Cela RG, Figueroa-Angulo EE (2010) Gene expression in trypanosomatid parasites. J Biomed Biotech 2010:525241

Marz M, Vanzo N, Stadler PF (2010) Temperature-dependent structural variability of RNAs: spliced leader RNAs and their evolutionary history. J Bioinform Comput Biol 8:1–17

Panigrahi AK, Ernst NL, Domingo GJ, Fleck M, Salavati R, Stuart KD (2006) Compositionally and functionally distinct editosomes in *Trypanosoma brucei*. RNA 2:1038–1049

Penny D, Hoeppner MP, Poole AM, Jeffares DC (2009) An overview of the introns-first theory. J Mol Evol 69:527–540

Poole AM, Logan DT (2005) Modern mRNA proofreading and repair: clues that the last universal common ancestor possessed an RNA genome? Mol Biol Evol 22:1444–1455

Poole AM, Jeffares D, Penny D (1998) The path from RNA world. J Mol Evol 46:1–17

Poole AM, Jeffares D, Penny D (1999) Prokaryotes, the new kids on the block. BioEssays 21:880–889

Pyle AM, Fedorova O, Waldsich C (2007) Folding of group II introns: a model for large multidomain RNAs? Trends Biochem Sci 32: 251

Roitzsch M, Pyle AM (2009) The linear form of group II intron catalyzes efficient autocatalytic reverse splicing, establishing a potential for mobility. RNA 15:473–482

Sachs JL, Bull JJ (2005) Experimental evolution of conflict mediation between genomes. Proc Natl Acad Sci USA 102:390–395

Schmidt U, Podar M, Stahl U, Perlman PS (1996) Mutations of the two-nucleotide bulge of D5 of a group II intron block splicing in vitro and in vivo: phenotypes and suppressor mutations. RNA 2:1161–1172

Singh J, Padgett RA (2009) Rates of in situ transcription and splicing in large human genes. Nat Struct Mol Biol 16:1128–1133

Stoltzfus A (1999) On the possibility of constructive neutral evolution. J Mol Evol 49:169–181

Valadkhan S, Jaladat Y (2010) The spliceosomal proteome: at the heart of the largest cellular ribonucleoprotein machine. Proteomics 10:4128–4141

Valentine DL (2007) Adaptations to energy stress dictate the ecology and evolution of the Archaea. Nat Rev Microbiol 5:316–323

Vesteg M, Krajčovič J (2008) Origin of eukaryotic cells as a symbiosis of parasitic α-proteobacteria in the periplasm of two-membrane-bounded sexual pre-karyotes. Commun Integr Biol 1:104–113

Vesteg M, Krajčovič J (2011) The falsifiability of the models for the origin of eukaryotes. Curr Genet 57:367–390

Vesteg M, Vacula R, Burey S, Löffelhardt W, Drahovská H, Martin W, Krajčovič J (2009) Expression of nucleus-encoded genes for chloroplast proteins in the flagellate *Euglena gracilis*. J Eukaryot Microbiol 56:159–166

Vesteg M, Vacula R, Steiner JM, Mateášiková B, Löffelhardt W, Brejová B, Krajčovič J (2010) A possible role for short introns in the acquisition of stroma-targeting peptides in the flagellate *Euglena gracilis*. DNA Res 17:223–231

Wernegreen JJ (2005) For better or worse: genomic consequences of intracellular mutualism and parasitism. Curr Opin Genet Dev 15:572–583

Woese CR (1998) The universal ancestor. Proc Natl Acad Sci USA 95:6854–6859

Woese CR (2000) Interpreting the universal phylogenetic tree. Proc Natl Acad Sci USA 97:8392–8396

Woese CR (2001) Translation: in retrospect and prospect. RNA 7:1055–1067

Woese CR (2002) On the evolution of cells. Proc Natl Acad Sci USA 99:8742–8747

Woolfit M, Bromham L (2003) Increased rates of sequence evolution in endosymbiotic bacteria and fungi with small effective population sizes. Mol Biol Evol 20:1545–1555