

REVIEW

Open Access



Transposon-derived introns as an element shaping the structure of eukaryotic genomes

Weronika Mikina¹, Paweł Hatakc¹ and Rafał Milanowski^{1*}

Abstract

The widely accepted hypothesis postulates that the first spliceosomal introns originated from group II self-splicing introns. However, it is evident that not all spliceosomal introns in the nuclear genes of modern eukaryotes are inherited through vertical transfer of intronic sequences. Several phenomena contribute to the formation of new introns but their most common origin seems to be the insertion of transposable elements. Recent analyses have highlighted instances of mass gains of new introns from transposable elements. These events often coincide with an increase or change in the spliceosome's tolerance to splicing signals, including the acceptance of noncanonical borders. Widespread acquisitions of transposon-derived introns occur across diverse evolutionary lineages, indicating convergent processes. These events, though independent, likely require a similar set of conditions. These conditions include the presence of transposon elements with features enabling their removal at the RNA level as introns and/or the existence of a splicing mechanism capable of excising unusual sequences that would otherwise not be recognized as introns by standard splicing machinery. Herein we summarize those mechanisms across different eukaryotic lineages.

Keywords Intron, Splicing, Evolution, Transposon, Transposable element, Introner, Introner element, MITE

Introduction

Introns are noncoding gene sequences that are excised during RNA maturation. Following their removal from the pre-RNA, a mature transcript is generated that includes only the exons. This mature transcript is the final RNA molecule that performs a specific function within the cell, such as serving as tRNA or rRNA, or is utilized by the cellular machinery in the translation process (mRNA) [1, 2]. Introns were initially characterized in the 1970s as "intervening sequences" that interrupt the coding DNA of genes [2]. Subsequently, exons were widely held to be the only important part of the transcript while introns should be rapidly excised and degraded, deemed

as junk sequences unnecessary for cellular function. However, contemporary research has provided ample evidence to refute this notion, demonstrating that introns may indeed perform vital functions. Introns have been implicated in nearly all stages of mRNA maturation, with their most significant roles encompassing alternative splicing and the regulation of gene expression. Additionally, certain introns encode functional RNA molecules, such as short interfering RNAs (siRNAs) and microRNAs (miRNAs) which interact with mature transcripts, leading to translational repression. Furthermore, introns influence numerous other cellular processes, including mRNA transport and RNA quality control [3, 4].

Introns are eliminated from precursor RNA through splicing, in which intron–exon junctions, known as splice sites, are cleaved, leading to the removal of the intron and the joining of exons to form a continuous RNA molecule. Depending on the mechanism of removal, several classes of introns are distinguished. Group I, II, and III introns comprise RNA molecules with autocatalytic properties

*Correspondence:

Rafał Milanowski
r.milanowski@uw.edu.pl

¹ Institute of Evolutionary Biology, Faculty of Biology, Biological and Chemical Research Centre, University of Warsaw, Żwirki i Wigury 101, Warsaw 02-089, Poland



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

and a conserved secondary and tertiary structure, enabling them to be excised from precursor RNA. These introns exhibit characteristics akin to mobile genetic elements capable of relocation within the genome [5, 6]. Conversely, tRNA introns are mostly found in tRNA genes of eukaryotes and Archaea (in certain cases, introns of this type also occur in pre-rRNA and pre-mRNA molecules). The excision of tRNA introns is a multistage process involving various enzymes, e.g. endonucleases and RNA ligases [7].

The most numerous and extensively studied group of introns are spliceosomal introns, intervening sequences unique to and nearly ubiquitous in eukaryotes [8]. The excision of introns of this type is facilitated by the spliceosome — a ribonucleoprotein complex consisting of RNA molecules and proteins. The pivotal components of the spliceosome are five small nuclear ribonucleoproteins (snRNPs) rich in uridine residues. Furthermore, the spliceosome comprises more than 200 protein splicing factors [9]. Spliceosomal introns are distinguished by the presence of evolutionarily conserved sequences at both the 5' and 3' ends, which are critical for accurate recognition and subsequent excision. These introns typically feature a GT dinucleotide (less frequently GC) at the 5' donor site and an AG dinucleotide at the 3' acceptor site. Additionally, there are other conserved motifs crucial for RNA splicing, such as the branchpoint sequence (usually an adenosine nucleotide) and the polypyrimidine tract located near the 3' end of the intron (Fig. 1). However, the exact position of the branchpoint site, as well as the position, length, and variability of the polypyrimidine tract (if present) differ among various eukaryotic lineages [10].

In addition to spliceosomal introns with canonical GT-AG junctions, many eukaryotic organisms harbor introns with noncanonical splice sites [1, 11, 12]. The largest subgroup of such introns comprises U12 introns which may feature unconventional junctions such as AT-AC. Despite the presence of noncanonical nucleotides at their termini, these introns are recognized and excised by the spliceosome. This is facilitated by the existence of another type of spliceosome in eukaryotic cells. While the U2 spliceosome (major spliceosome) is primarily responsible for excising the majority of spliceosomal introns with canonical GT-AG junctions, the U12 spliceosome (minor spliceosome) specializes in removing introns with specific features beyond the junction sequences themselves. While most of these introns have canonical GT-AG borders, a substantial portion of them possess AT-AC junctions. Unlike the U2 introns, U12 introns lack a distinct polypyrimidine tract. Moreover, the sequences surrounding the 5' end and the branchpoint are longer and more conserved than those found in U2 introns [13]. U12 introns are found across most

eukaryotes and typically constitute less than 0.5% of all introns in the genome [9, 14].

The origin of spliceosomal introns

The presence of introns in the genome entails increased energy expenditure for the cell. The splicing process involves one of the largest molecular complexes in the cell and is energetically costly and time-consuming [3]. Moreover, a potential mutation at the intron/exon boundary can be particularly detrimental to the organism as it hinders junction recognition and intron excision, potentially resulting in the production of a nonfunctional gene product. Therefore investigating the origin and functions of introns has become the subject of much consideration. One hypothesis, for instance, suggests that introns emerged in genomes as selfish sequences, replicating themselves at the host's expense and acquiring numerous functions only later in various evolutionary lineages [3].

Two hypotheses have been proposed to explain the origin and accumulation of introns in genomes. The first hypothesis, known as "introns-early", suggests that introns were present in the early stages of genome evolution. According to this view, the earliest genes contained numerous introns that played a crucial role in genome reorganization and the generation of new proteins by facilitating the recombination of sequences encoding individual polypeptides. The second hypothesis, termed "introns-late", proposes that introns evolved exclusively in eukaryotes and have since accumulated in genomes [1, 11]. Despite these considerations, there is rather no doubt that the LECA genome had numerous introns [15, 16]. Another hypothesis worth mentioning suggests a significant role of endosymbiosis in the emergence of introns in Eukarya. The acquisition of mitochondria through endosymbiosis with α -proteobacteria would initiate the transfer of self-splicing group II introns to the host genome where they evolved into spliceosomal introns over time [1]. The similarity between molecular mechanisms of the excision of these introns and group II introns, along with genome reconstructions of the last eukaryotic common ancestor (LECA), support this hypothesis [1].

Although the hypothesis that spliceosomal introns and small nuclear RNAs (snRNAs) originated from group II introns is widely accepted, it is evident that not all spliceosomal introns in the nuclear genes of modern eukaryotes stem from intronic sequences inherited through vertical transfer. Several phenomena have been described that could contribute to the formation of new introns in the genome. These include the acquisition of an intron during DNA repair following a double-strand break, or intronization of an exon sequence [17, 18]. The creation

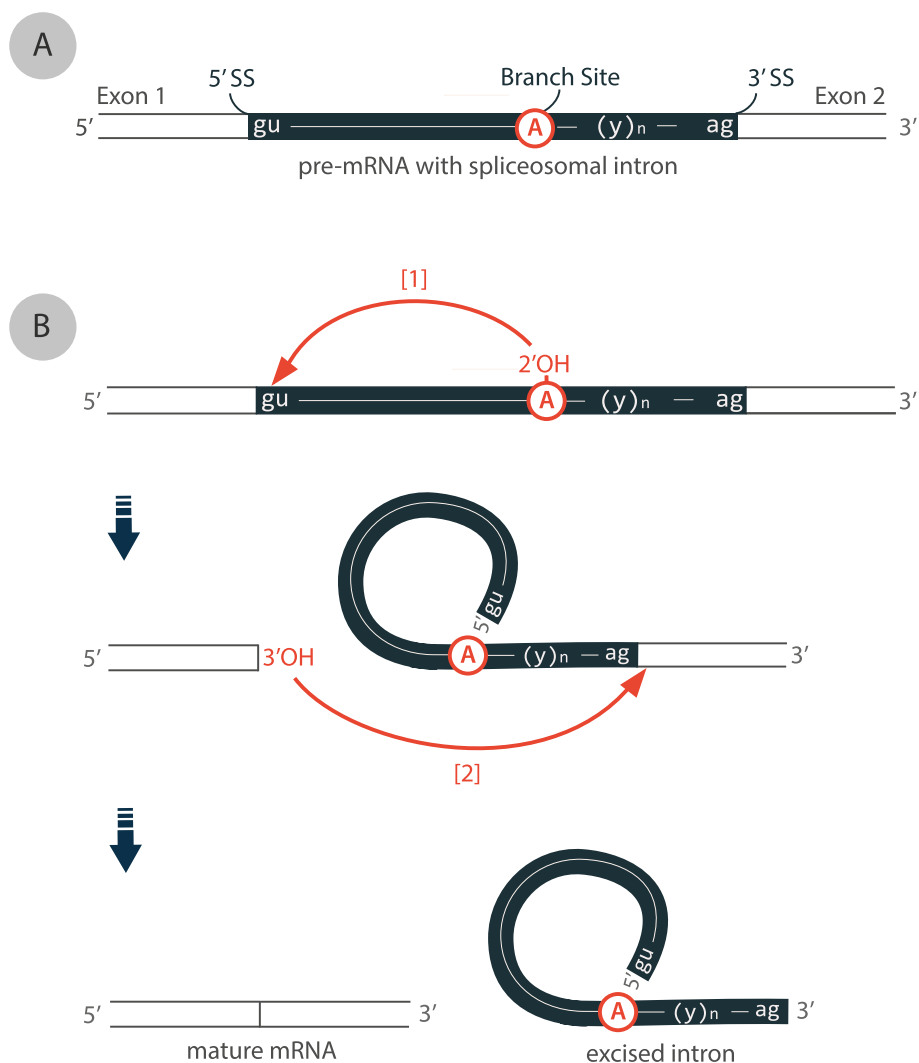


Fig. 1 Conserved nucleotides **(A)** and schematic representation of spliceosomal intron excision **(B)**. Branchpoint, poly(pyrimidine) tract, as well as both 5' and 3' splice sites are indicated. All of them play a key role in the process of removing spliceosomal introns, which consist of two sequential transesterification reactions. (1) The 2' hydroxyl (2'OH) group of a branch site performs a nucleophilic attack on the 5' splice site, leading to the formation of the lariat intermediate. (2) The released 3'OH group from the exon conducts a nucleophilic attack on the 3' splice site, leading to exons joining and excision of the intron lariat

of new introns from nonintronic sequences requires the presence of signals recognizable by the cell's splicing machinery, enabling their efficient removal from pre-mRNA.

Spliceosomal introns derived from transposable elements

Recent genomic analyses have highlighted the frequent gain of new introns through the insertion of sequences originating from transposons [19–22]. Depending on the site where transposons are integrated within the genome, transposon insertions can significantly impact organismal function. While insertions into non-coding

sequences can generally be considered relatively benign, insertion within an exon typically leads to harmful changes, usually resulting in gene inactivation [23]. However, if sequences resembling spliceosome recognition signals are present within the transposon or in the vicinity of the insertion site, the insertion can lead to a transformation of the transposon into an intron allowing their safe integration [24, 25]. These new introns often exhibit short duplicated sequences (target site duplications, TSDs) at exon–intron junctions or in their immediate vicinity as a result of repairing staggered single-stranded regions generated by target DNA cleavage. Additionally,

terminal inverted repeats (TIRs) are frequently present, confirming their transposon origin (Fig. 2).

Singular introns of transposon or retrotransposon origin have been reported in many eukaryotes [20–22]. Nevertheless, instances of mass gain of new introns from transposon sequences, which significantly reshaped entire genomes, are particularly intriguing. At the outset, thousands of introns with similar sequences were observed in the nuclear genome of the green alga *Micromonas pusilla* [26]. These repetitive intronic sequences were termed introner elements (IEs), although their association with transposable elements was not immediately recognized. Subsequent analysis revealed that these sequences are flanked by three-nucleotide target site duplications (TSDs) and exhibit terminal inverted repeats (TIRs) adjacent to the duplicated sequences. Consequently, introners were redefined as short DNA transposons, similar to MITE transposons (miniature inverted-repeat transposable elements) which are nonautonomous transposon elements that have been described extensively in various plant taxa as well as in the genomes of animals, fungi, and bacteria [27]. At the RNA level, each introner in

M. pusilla is predisposed to be removed, serving as a donor of one splice site, with the second splice site created within the duplicated target sequence (Fig. 2). This facilitates the effective removal of the introner from the mRNA [28]. Subsequently, introners and introner-like elements (ILEs) were discovered in various organisms, including fungi [29, 30], stramenopiles [28], and dinoflagellates [31, 32]. An extensive analysis of genomic sequences for introners revealed their prevalence, with 5.2% of genomes containing IE-derived introns [33]. Over 27,000 such introns were categorized into 548 families based on similarity. It was also noted that the vast majority of introners seem to be DNA transposons. Notably, not all analyzed introners exhibit TSD and TIR sequences. Splicing signals enabling excision of IE-derived introns may originate from the introner sequence itself, TSDs, or exon sequences. The presence of IE-derived introns spans eight independent phylogenetic lineages across six major evolutionary groups of eukaryotes. Moreover, organisms inhabiting aquatic environments are 6.5 times more likely to contain introners, suggesting the possibility of their spreading via horizontal transfer in aquatic environments [33].

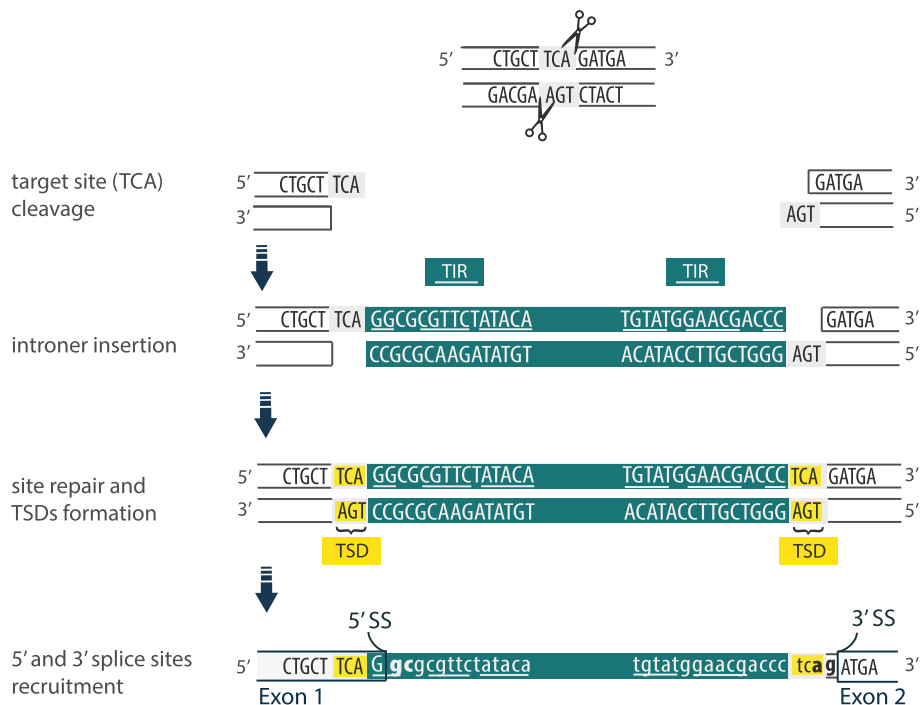


Fig. 2 Introner insertion leading to the formation of an intron. After transposase cuts the TCA site, introner (marked in green) integrates into the target sequence. Site repair results in the duplication of this sequence on each side of the introner. These short, duplicated sequences (TSD) are highlighted in yellow. Additionally, the terminal inverted repeats (TIR) are underlined. Restored splicing sites (bold) of newly gained intron (lowercase) are either carried by introner (5'), or co-opted from TSD (3'). The sequence originates from *M. pusilla* and illustrates the potential sequence of events leading to the IE-derived intron gain [17, 28]

Noncanonical introns and the evolution of the spliceosome

Widespread acquisition of new introns from transposons often coincides with an increase or change in the tolerance of the spliceosome to splicing signals, including the acceptance of noncanonical borders (other than GT-AG). A significant proportion of introns with noncanonical junctions have been identified in the genome of the tunicate *Fritillaria borealis* [34]. Genomic and transcriptomic analyses revealed that AG-AC and AG-AT are the most prevalent splice sites in this organism, although various other dinucleotide combinations have also been observed. Introns terminating in GT-AG in *F. borealis* are relatively rare and typically occupy evolutionarily conserved positions within genes. Conversely, noncanonical introns are found at species-specific locations, suggesting recent acquisition. Furthermore, *F. borealis* introns exhibit TIR inverted repeats and TSD sequences, indicative of their transposon origin (Fig. 3A).

Introns with noncanonical junctions in the *F. borealis* genome likely originate, as suggested for introners in *Micromonas*, from MITEs. However, the mechanism underlying the excision of these introns is particularly intriguing. It appears that tunicate introns are efficiently excised by the spliceosome despite noncanonical ends, but only in *F. borealis*. In contrast, the same introns are not spliced out by the human spliceosome. This discrepancy suggests an evolutionary change in the spliceosome of tunicates, enabling the neutralization of the effects of transposon insertions by adapting to the removal of introns with noncanonical junctions [34].

Adaptation of the spliceosome to remove unusual introns has also been observed in marine parasitic

dinoflagellates of the genus *Amoebophyra* [31]. Protists within this group are characterized by large genomes with unusual organization. Two strains of *Amoebophyra* exhibit significant variability in intron boundaries, with more than 60% of introns being noncanonical. These atypical introns differ in length and GC content compared to their canonical counterparts, exist in multiple copies and exhibit TSDs and TIRs, indicating their transposon origin (Fig. 3B). Additionally, in both tested strains, the U1 snRNA—which is crucial for recognizing the donor end of the intron—was apparently absent. The lack of U1 snRNA suggests the development of an alternative splicing mechanism, possibly involving recruitment of a new subunit to the spliceosome complex, facilitating effective removal of introns with unusual borders [31].

An increased number of unusual intron boundaries has also been observed in other dinoflagellates, such as *Symbiodinium* species and *Polarella glacialis* [32]. The genomes of these protists harbor a substantial number of newly acquired IE-derived introns, while ancient introns have undergone extensive loss. Analysis of IE introns revealed a prevalence of introns with GC-AG borders, a feature that typically occurs on a small scale in all eukaryotes but is dominant in the aforementioned dinoflagellates (Fig. 3C). In this context, the reduced/changed selectivity of the spliceosome may facilitate the acquisition of new introns due to the proliferation of introners, generating atypical splicing signals. The authors propose a model to elucidate the evolution of such unusual introns. According to this model, the initial stage involves a massive loss of ancient, canonical introns. A decreased number of introns with highly homogeneous junctions

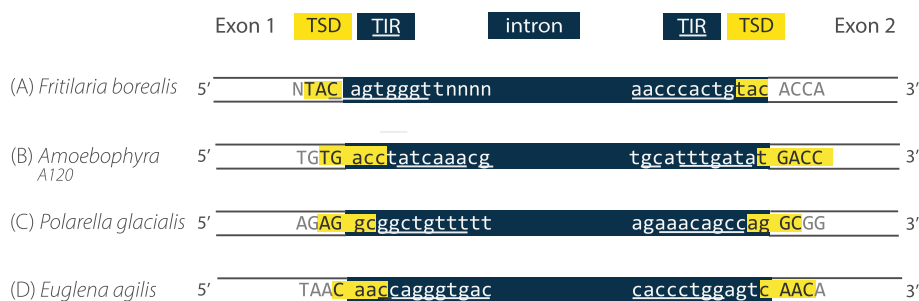


Fig. 3 Diversity of transposon-derived introns with noncanonical junctions. Most of them exhibit specific sequence features: TIR (underlined) and TSD at the intron/exon boundaries (highlighted in yellow), suggesting their origin; intron sequences in lowercase. **A** In the genome of the tunicate, *F. borealis*, the most prevalent splice sites in this organism are AG-AC and AG-AT. The majority of noncanonical introns are typically preceded by an exonic sequence TAC that led to the formation of the 3' splice site [34]. **B** The TSDs motifs of 3–5 nt length in *Amoebophyra* show high variability, whereas some of the TIR motifs were conserved and strain-specific [31]. **C** Introns in *P. glacialis* demonstrate significant diversity, and have been categorized into 15 separate families, revealing various patterns of intron acquisition and the recognition of new splicing sites [32]. **D** Nonconventional introns in euglenids form a stable secondary structure based on TIR (usually CAG and CTG in positions +4, 5, 6 and -8, 7, 6 respectively) nucleotides base-pairing, that brings the ends of the intron together. They exhibit noncanonical junctions, with often repetitions of TSD at the intron/exon boundaries [36, 37]

may lead to decreased selectivity for recognized splicing signals. Less constrained spliceosomes may have a greater capacity to adapt to gene-disrupting transposable elements, increasing the probability of the emergence of transposon-derived introns, even in cases of noncanonical junctions [32].

The relationship described above between the massive gain of introns from transposon elements and the presence of atypical splicing signals extends beyond spliceosomal U2 introns. Analysis of the *Physarum polycephalum* genome revealed a significant abundance of U12 introns, approximately 25 times greater than that of any other species [35]. Furthermore, these introns with atypical borders appear to be a relatively recent acquisition. They frequently exhibit short inverted repeats at their ends, suggesting a possible transposon origin. Interestingly, unlike those in other species, new U12 introns demonstrate high removal efficiency from transcripts, likely facilitating a substantial increase in their number [35].

Another intriguing case of unusual intron spread is observed in euglenids. While these organisms possess typical spliceosomal U2 introns, they also harbor a distinct group of nonconventional introns with markedly different characteristics [36, 37]. These nonconventional introns feature noncanonical junctions, often consistent with the pyrimidine|purine consensus on both ends, lack a polypyrimidine tract but exhibit TIRs which bring both ends of the intron closer together at the RNA level. Additionally, repeats resembling target site duplications (TSD-like sequences) are frequently found at intron–exon junctions (Fig. 3D). Moreover, their insertions into new positions within genes has been observed, further suggesting their transposon origin, likely from MITE elements. The length of nonconventional introns varies widely, ranging from several dozen to several thousand nucleotides, with no clear pattern within the group. The RNA secondary structure of these introns is somewhat conserved, particularly near the ends [36, 37]. The removal of nonconventional introns occurs after the removal of spliceosomal conventional introns, and upon excision, they manifest in the cell as circular RNA molecules with full-length ends lacking the typical lariat form [38, 39]. The mechanism underlying the removal of these introns from transcripts has not been elucidated. However, discernible differences suggest that their removal involves a nonconventional, additional spliceosome or a spliceosome-independent process. This observation indicates the development of a novel splicing mechanism that likely facilitates the acquisition of new intron sequences of transposon origin. Thus, this phenomenon seems to be a scenario similar to that of previously discussed transposon-derived U2 introns, in which the presence

of a spliceosome with a greater or different tolerance for splicing signals promotes the emergence of new intervening sequences.

Summary

Recent research indicates that the genomes of many eukaryotic organisms undergo constant changes. While instances of intron loss are observed, there are also processes of mass intron gain. This phenomenon, while increasing the costs associated with maintaining and expressing the genome, simultaneously shapes its structure, enhancing the flexibility of gene expression and expanding the repertoire of available proteins within the cell. Widespread gains of transposon-derived introns are observed across diverse evolutionary lineages, indicating convergent processes. These events occur independently but likely result from common conditions: the presence of transposon elements with features enabling their removal at the RNA level and/or the existence of a splicing mechanism capable of excising unusual introns that would otherwise not be recognized by standard mechanisms. Our expanding understanding of the dynamics of intron loss and gain not only sheds light on the evolution of eukaryotic genomes but also provides insights into the evolutionary processes that gave rise to spliceosomal introns and the complex splicing machinery in LECA.

Acknowledgements

Not applicable

Authors' contributions

WM wrote the draft, RM the final version of manuscript. WM prepared figures, PH inspected the manuscript critically. WM, PH and RM took part in the revision of manuscript. All authors read and approved the final manuscript.

Funding

BOB-IDUB-622–325/2022 IDUB New Ideas 2B in Priority Research Area I.

Availability of data and materials

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 30 January 2024 Accepted: 23 July 2024

Published online: 27 July 2024

References

1. Rogozin IB, Carmel L, Csuros M, Koonin EV. Origin and evolution of spliceosomal introns. *Biol Direct*. 2012;7:11.

2. Gilbert W. Why genes in pieces? *Nature*. 1978;271:501–501.
3. Chorev M, Carmel L. The function of introns. *Front Genet*. 2012;3:55.
4. Girardini KN, Olthof AM, Kanadia RN. Introns: the “dark matter” of the eukaryotic genome. *Front Genet*. 2023;14:1150212.
5. Gogarten JP, Hilario E. Inteins, introns, and homing endonucleases: recent revelations about the life cycle of parasitic genetic elements. *BMC Evol Biol*. 2006;6:94.
6. Hausner G, Hafez M, Edgell DR. Bacterial group I introns: mobile RNA catalysts. *Mob DNA*. 2014;5:8.
7. Hopper AK, Nostramo RT. tRNA processing and subcellular trafficking proteins multitask in pathways for other RNAs. *Front Genet*. 2019;10:96.
8. William Roy S, Gilbert W. The evolution of spliceosomal introns: patterns, puzzles and progress. *Nat Rev Genet*. 2006;7:211–21.
9. Will CL, Luhrmann R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol*. 2011;3:a003707.
10. Irimia M, Roy SW. Evolutionary convergence on highly-conserved intron structures in intron-poor eukaryotes and insights into the ancestral eukaryotic genome. *PLoS Genet*. 2008;4:e1000148.
11. Poverennaya IV, Roytberg MA. Spliceosomal introns: features, functions, and evolution. *Biochem Mosc*. 2020;85:725–34.
12. Frey K, Pucker B. Animal, fungi, and plant genome sequences harbor different non-canonical splice sites. *Cells*. 2020;9:458.
13. Moyer DC, Larue GE, Hershberger CE, Roy SW, Padgett RA. Comprehensive database and evolutionary dynamics of U12-type introns. *Nucleic Acids Res*. 2020;48:7066–78.
14. Lin CF, Mount SM, Jarmołowski A, Makalowski W. Evolutionary dynamics of U12-type spliceosomal introns. *BMC Evol Biol*. 2010;10:47.
15. Csuros M, Rogozin IB, Koonin EV. A detailed history of intron-rich eukaryotic ancestors inferred from a global survey of 100 complete genomes. *PLoS Comput Biol*. 2011;7:e1002150.
16. Koonin EV. Intron-dominated genomes of early ancestors of Eukaryotes. *J Hered*. 2009;100:618–23.
17. Yenerall P, Zhou L. Identifying the mechanisms of intron gain: progress and trends. *Biol Direct*. 2012;7:29.
18. Ignatenko A, Gumińska N, Milanowski R. Mechanizmy utraty i nabywania intronów spliceosomalnych. *Postepy Biochem*. 2019;65:289–98.
19. Sela N, Mersch B, Gal-Mark N, Lev-Maor G, Hotz-Wagenblatt A, Ast G. Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu’s unique role in shaping the human transcriptome. *Genome Biol*. 2007;8:R127.
20. Denoeud F, Henriët S, Mungpakdee S, Aury JM, Da SC, Brinkmann H, Mikhaleva J, Olsen LC, Jubin C, Canestro C, Bouquet JM, Danks G, Poulain J, Campsteijn C, Adamski M, Cross I, Yadetie F, Muffato M, Louis A, Butcher S, Tsagkogeorga G, Konrad A, Singh S, Jensen MF, Cong EH, Eikeseth-Otteraa H, Noel B, Anthouard V, Porcel BM, Kachouri-Lafond R, Nishino A, Ugolini M, Chourrout P, Nishida H, Aasland R, Huzurbazar S, Westhof E, Delsuc F, Levrach H, Reinhardt R, Weissenbach J, Roy SW, Artiguenave F, Postlethwait JH, Manak JR, Thompson EM, Jaillon O, Du PL, Boudinot P, Liberles DA, Volf JN, Philippe H, Lenhard B, Roest CH, Wincker P, Chourrout D. Plasticity of animal genome architecture unmasked by rapid evolution of a pelagic tunicate. *Science*. 2010;330:1381–5.
21. Yenerall P, Krupa B, Zhou L. Mechanisms of intron gain and loss in *Drosophila*. *BMC Evol Biol*. 2011;11:364.
22. Wu B, Macielog AI, Hao W. Origin and spread of spliceosomal introns: insights from the fungal clade zygomycota. *Genome Biol Evol*. 2017;9:1658–2667.
23. Ayarpadikannan S, Kim HS. The impact of transposable elements in genome evolution and genetic instability and their implications in various diseases. *Genomics Inform*. 2014;12:98.
24. Iwamoto M, Maekawa M, Saito A, Higo H, Higo K. Evolutionary relationship of plant catalase genes inferred from exon-intron structures: isozyme divergence after the separation of monocots and dicots. *Theor Appl Genet*. 1998;97:9–19.
25. Roy SW. The origin of recent introns: transposons? *Genome Biol*. 2004;5:251.
26. Worden AZ, Lee JH, Mock T, Rouzé P, Simmons MP, Aerts AL, et al. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science*. 2009;324:268–72.
27. Fattash I, Rooke R, Wong A, Hui C, Luu T, Bhardwaj P, et al. Miniature inverted-repeat transposable elements: discovery, distribution, and activity. *Genome*. 2013;56:475–86.
28. Huff JT, Zilberman D, Roy SW. Mechanism for DNA transposons to generate introns on genomic scales. *Nature*. 2016;538:533–6.
29. van der Burgt A, Severing E, de Wit PJGM, Collemare J. Birth of new spliceosomal introns in fungi by multiplication of introner-like elements. *Curr Biol*. 2012;22:1260–5.
30. Collemare J, van der Burgt A, de Wit PJGM. At the origin of spliceosomal introns: Is multiplication of introner-like elements the main mechanism of intron gain in fungi? *Commun Integr Biol*. 2013;6:e23147.
31. Farhat S, Le P, Kayal E, Noel B, Bigeard E, Corre E, et al. Rapid protein evolution, organellar reductions, and invasive intronic elements in the marine aerobic parasite dinoflagellate *Amoebophrya* spp. *BMC Biol*. 2021;19:1.
32. Roy SW, Gozashti L, Bowser BA, Weinstein BN, Larue GE, Corbett-Detig R. Intron-rich dinoflagellate genomes driven by introner transposable elements of unprecedented diversity. *Curr Biol*. 2023;33:189–96.
33. Gozashti L, Roy SW, Thornlow B, Kramer A, Ares M, Corbett-Detig R. Transposable elements drive intron gain in diverse eukaryotes. *Proc Natl Acad Sci USA*. 2022;119:e2209766119.
34. Henriët S, Colom Sanmartí B, Sumic S, Chourrout D. Evolution of the U2 spliceosome for processing numerous and highly diverse non-canonical introns in the chordate *fritillaria borealis*. *Curr Biol*. 2019;29:3193–9.
35. Larue GE, Eliáš M, Roy SW. Expansion and transformation of the minor spliceosomal system in the slime mold *Physarum polycephalum*. *Curr Biol*. 2021;31:3125–31.
36. Milanowski R, Gumińska N, Karnkowska A, Ishikawa T, Zakryś B. Intermediate introns in nuclear genes of euglenids – are they a distinct type? *BMC Evol Biol*. 2016;16:49.
37. Milanowski R, Karnkowska A, Ishikawa T, Zakryś B. Distribution of conventional and nonconventional introns in tubulin (α and β) genes of euglenids. *Mol Biol Evol*. 2014;31:584–93.
38. Gumińska N, Plecha M, Zakryś B, Milanowski R. Order of removal of conventional and nonconventional introns from nuclear transcripts of *Euglena gracilis*. *PLoS Genet*. 2018;14:e1007761.
39. Gumińska N, Zakryś B, Milanowski R. A new type of circular RNA derived from nonconventional introns in nuclear genes of Euglenids. *J Mol Biol*. 2021;433:166758.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.