



Peer Community In Evolutionary Biology

The drift barrier hypothesis and the limits to alternative splicing accuracy

Ignacio Bravo  based on peer reviews by **Lars M. Jakt** and 2 anonymous reviewers

Florian Benitiere, Anamaria Necsulea, Laurent Duret (2023) Random genetic drift sets an upper limit on mRNA splicing accuracy in metazoans. bioRxiv, ver. 4, peer-reviewed and recommended by Peer Community in Evolutionary Biology.

<https://doi.org/10.1101/2022.12.09.519597>

Submitted: 12 December 2022, Recommended: 25 September 2023

Cite this recommendation as:

Bravo, I. (2023) The drift barrier hypothesis and the limits to alternative splicing accuracy. *Peer Community in Evolutionary Biology*, 100642. [10.24072/pci.evolbiol.100642](https://doi.org/10.24072/pci.evolbiol.100642)

Published: 25 September 2023

Copyright: This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Accurate information flow is central to living systems. The continuity of genomes through generations as well as the reproducible functioning and survival of the individual organisms require a faithful information transfer during replication, transcription and translation. The differential efficiency of natural selection against “mistakes” results in decreasing fidelity rates for replication, transcription and translation. At each level in the information flow chain (replication, transcription, translation), numerous complex molecular systems have evolved and been selected for preventing, identifying and, when possible, correcting or removing such “mistakes” arising during information transfer.

However, fidelity cannot be improved *ad infinitum*. First, because of the limits imposed by the physical nature of the processes of copying and recoding information over different molecular supports: all mechanisms ensuring fidelity during biological information transfer ultimately rely on chemical kinetics and thermodynamics. The more accurate a copying process is, the lower the synthesis rate and the higher the energetic cost of correcting errors. Second, because of the limits imposed by random genetic drift: natural selection cannot effectively act on an allele that contributes with a small differential advantage unless effective population size is large. If $s < 1/Ne$ (or $s < 1/(2Ne)$ in diploids) the allele frequency in the population is *de facto* subject to neutral drift processes.

In their preprint “*Random genetic drift sets an upper limit on mRNA splicing accuracy in metazoans*”, Bénitière, Necsulea and Duret explore the validity of this last mentioned “drift barrier” hypothesis for the case study of alternative splicing diversity in eukaryotes (Bénitière et al. 2022). Splicing refers to an ensemble of eukaryotic molecular processes mediated by a large number of proteins and ribonucleoproteins and involving nucleotide sequence recognition, that uses as a molecular substrate a precursor messenger RNA (mRNA), directly transcribed from the DNA, and produces a mature mRNA by removing introns and joining exons (Chow et al. 1977).

Alternative splicing refers to the case in which different molecular species of mature mRNAs can be produced, either by *cis*-splicing processes acting on the same precursor mRNA, e.g. by varying the presence/absence of different exons or by varying the exon-exon boundaries, or by *trans*-splicing processes, joining exons from different precursor mRNA molecules.

The diversity of mRNA molecular species generated by alternative splicing enlarges the molecular phenotypic space that can be generated from the same genotype. In humans, alternative splicing occurs in around 95% of the *ca.* 20,000 genes, resulting in *ca.* 100,000 medium-to-high abundance transcripts (Pan et al. 2008). In multicellular organisms, the frequency of alternatively spliced mRNAs varies between tissues and across ontogeny, often in a switch-like pattern (Wang et al. 2008). In the molecular and cell biology community, it is commonly accepted that splice variants contribute with specific functions (Marasco and Kornblihtt 2023) although there exists a discussion around the functional nature of low-frequency splice variants (see for instance the debate between Tress et al. 2017 and Blencowe 2017). The origin, diversity, regulation and evolutionary advantage of alternative splicing constitutes thus a playground of the selectionist-neutralist debate, with one extreme considering that most splice variants are mere “mistakes” of the splicing process (Pickrell et al. 2010), and the other extreme considering that alternative splicing is at the core of complexity in multicellular organisms, as it increases the genome coding potential and allows for a large repertoire of cell types (Chen et al. 2014).

In their manuscript, Bénétière, Necsulea and Duret set the cursor towards the neutralist end of the gradient and test the hypothesis of whether the high alternative splice rate in “complex” organisms corresponds to a high rate of splicing “mistakes”, arising from the limit imposed by the drift barrier effect on the power of natural selection to increase accuracy (Bush et al. 2017). In their preprint, the authors convincingly show that in metazoans a fraction of the variation of alternative splicing rate is explained by variation in proxies of population size, so that species with smaller *Ne* display higher alternative splice rates. They communicate further that abundant splice variants tend to preserve the reading frame more often than low-frequency splice variants, and that the nucleotide splice signals in abundant splice variants display stronger evidence of purifying selection than those in low-frequency splice variants. From all the evidence presented in the manuscript, the authors interpret that “*variation in alternative splicing rate is entirely driven by variation in the efficacy of selection against splicing errors*”.

The authors honestly present some of the limitations of the data used for the analyses, regarding i) the quality of the proxies used for *Ne* (i.e. body length, longevity and dN/dS ratio); ii) the heterogeneous nature of the RNA sequencing datasets (full organisms, organs or tissues; different life stages, sexes or conditions); and iii) mostly short RNA reads that do not fully span individual introns. Further, data from bacteria do not verify the herein communicated trends, as it has been shown that bacterial species with low population sizes do not display higher transcription error rates (Traverse and Ochman 2016). Finally, it will be extremely interesting to introduce a larger evolutionary perspective on alternative splicing rates encompassing unicellular eukaryotes, in which an intriguing interplay between alternative splicing and gene duplication has been communicated (Hurtig et al. 2020).

The manuscript from Bénétière, Necsulea and Duret makes a significant advance to our understanding of the diversity, the origin and the physiology of post-transcriptional and post-translational mechanisms by emphasising the fundamental role of non-adaptive evolutionary processes and the upper limits to splicing accuracy set by genetic drift.

References:

Bénétière F, Necsulea A, Duret L. 2023. Random genetic drift sets an upper limit on mRNA splicing accuracy in metazoans. bioRxiv, ver. 4 peer-reviewed and recommended by Peer Community in Evolutionary Biology. <https://doi.org/10.1101/2022.12.09.519597>

Blencowe BJ. 2017. The Relationship between Alternative Splicing and Proteomic Complexity. Trends Biochem Sci 42:407–408. <https://doi.org/10.1016/j.tibs.2017.04.001>

Bush SJ, Chen L, Tovar-Corona JM, Urrutia AO. 2017. Alternative splicing and the evolution of phenotypic novelty. *Philos Trans R Soc Lond B Biol Sci* 372:20150474.

<https://doi.org/10.1098/rstb.2015.0474>

Chen L, Bush SJ, Tovar-Corona JM, Castillo-Morales A, Urrutia AO. 2014. Correcting for differential transcript coverage reveals a strong relationship between alternative splicing and organism complexity. *Mol Biol Evol* 31:1402–1413. <https://doi.org/10.1093/molbev/msu083>

Chow LT, Gelinis RE, Broker TR, Roberts RJ. 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* 12:1–8. [https://doi.org/10.1016/0092-8674\(77\)90180-5](https://doi.org/10.1016/0092-8674(77)90180-5)

Hurtig JE, Kim M, Orlando-Coronel LJ, Ewan J, Foreman M, Notice L-A, Steiger MA, van Hoof A. 2020. Origin, conservation, and loss of alternative splicing events that diversify the proteome in *Saccharomyces* budding yeasts. *RNA* 26:1464–1480. <https://doi.org/10.1261/rna.075655.120>

Marasco LE, Kornblihtt AR. 2023. The physiology of alternative splicing. *Nat Rev Mol Cell Biol* 24:242–254. <https://doi.org/10.1038/s41580-022-00545-z>

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* 40:1413–1415.

<https://doi.org/10.1038/ng.259>

Pickrell JK, Pai AA, Gilad Y, Pritchard JK. 2010. Noisy splicing drives mRNA isoform diversity in human cells. *PLoS Genet* 6:e1001236. <https://doi.org/10.1371/journal.pgen.1001236>

Traverse CC, Ochman H. 2016. Conserved rates and patterns of transcription errors across bacterial growth states and lifestyles. *Proc Natl Acad Sci U S A* 113:3311–3316.

<https://doi.org/10.1073/pnas.1525329113>

Tress ML, Abascal F, Valencia A. 2017. Alternative Splicing May Not Be the Key to Proteome Complexity. *Trends Biochem Sci* 42:98–110. <https://doi.org/10.1016/j.tibs.2016.08.008>

Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* 456:470–476.

<https://doi.org/10.1038/nature07509>

Reviews

Evaluation round #3

Reviewed by anonymous reviewer 2, 05 September 2023

Please find my review attached.

[Download the review](#)

Evaluation round #2

DOI or URL of the preprint: <https://doi.org/10.1101/2022.12.09.519597>

Version of the preprint: 3

Authors' reply, 23 July 2023

[Download author's reply](#)

[Download tracked changes file](#)

Decision by [Ignacio Bravo](#) , posted 01 July 2023, validated 02 July 2023

Proposal for minor text and mathematical model revision.

The two reviewers have provided an extensive analysis of the revised version of the manuscript. Both of them agree that the manuscript has definitely improved in readability and that the analyses are now easier to follow and to understand, and I largely agree with them. The authors have also struggled to provide a complete description of the data and of the pipeline used to analyse them. Nevertheless, one of the reviewers points out a number of still ill-defined steps that may merit a proper description, for the sake of clarity but also for the interest of reproducibility and for the extension of the analyses to larger or finer datasets in the future. Further, one of the reviewers requires clarification about one of the main variables used in the analyses (i.e. "the definition of the AS rate of introns"), which is central to the analyses. Finally, one of the reviewers expresses once again their concerns regarding the value of the mathematical model included in the text, as they consider that the results obtained are directly derived from the assumptions and boundary conditions used to run the model. These points need to be clarified.

I am very excited by the results presented in this text, and I think it will make a significant contribution to the field, but that it may still require the above mentions points to be properly addressed.

Reviewed by anonymous reviewer 2, 03 May 2023

[Download the review](#)

Reviewed by [Lars M. Jakt](#), 16 June 2023

[Download the review](#)

Evaluation round #1

DOI or URL of the preprint: <https://doi.org/10.1101/2022.12.09.519597>

Version of the preprint: 2

Authors' reply, 07 April 2023

[Download author's reply](#)

[Download tracked changes file](#)

Decision by [Ignacio Bravo](#) , posted 13 February 2023, validated 13 February 2023

Recommendation for revision of the first version of the manuscript.

The preprint by Bénitière and coworkers has been evaluated by three experts in the field. The three reviewers agree on the importance and on the interest of the manuscript, in terms of relevance of the question addressed as well as in terms of validating a null, neutralist perspective. I largely share the enthusiasm of the the reviewers, which is very evident in their comments. Notwithstanding, I also share with the reviewers some concerns, concerning the concordance (or the lack of concordance) with previous works in the literature addressing similar or parallel questions around fidelity during transcription. Further, the existence and potential impact of

possible systematic biases linked to a diversity of sequencing technologies in the databases analysed, and to a differential sensitivity for detecting exon junctions needs to be more explicitly addressed in the text. The reviewers have extensively commented the text in several passages, and I globally agree that addressing the different concerns raised will undoubtedly ameliorate the manuscript for its clarity and soundness. I am convinced that an improved version may be a substantial contribution to the field.

Reviewed by anonymous reviewer 1, 15 January 2023

[Download the review](#)

Reviewed by Lars M. Jakt, 29 January 2023

[Download the review](#)

Reviewed by anonymous reviewer 2, 02 February 2023

Dear authors,
Please find my review in the attached pdf file.

[Download the review](#)