



# Small Open Reading Frames: How Important Are They for Molecular Evolution?

Diego Guerra-Almeida<sup>1\*</sup> and Rodrigo Nunes-da-Fonseca<sup>1,2\*</sup>

<sup>1</sup> Institute of Biodiversity and Sustainability, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil, <sup>2</sup> National Institute of Science and Technology in Molecular Entomology, Rio de Janeiro, Brazil

**Keywords:** small ORF, junk DNA, coding potential maturation, causal roles, selected effects, non-coding RNA, alternative ORF, pervasive translation

## OPEN ACCESS

### Edited by:

Fatemeh Maghuly,  
University of Natural Resources and  
Life Sciences Vienna, Austria

### Reviewed by:

Benedikt Obermayer,  
Charité Medical University of  
Berlin, Germany  
Mona Wu Orr,  
Amherst College, United States

### \*Correspondence:

Diego Guerra-Almeida  
diegoguerra@ufrj.br  
Rodrigo Nunes-da-Fonseca  
mfonseca@macae.ufrj.br

### Specialty section:

This article was submitted to  
Systems Biology,  
a section of the journal  
Frontiers in Genetics

**Received:** 21 June 2020

**Accepted:** 25 August 2020

**Published:** 20 October 2020

### Citation:

Guerra-Almeida D and  
Nunes-da-Fonseca R (2020) Small  
Open Reading Frames: How  
Important Are They for Molecular  
Evolution? *Front. Genet.* 11:574737.  
doi: 10.3389/fgene.2020.574737

## INTRODUCTION

Small Open Reading Frames (small ORFs/sORFs/smORFs) are important sources of putative peptides previously dismissed as being non-functional or junk DNA, as determined by early gene prediction methods. In fact, smORFs of <100 codons are possible coding sequences but sufficiently small to occur very frequently and randomly in genomes; thus, the detection of their coding potential and functional assessment is similar to a walk in the dark. Furthermore, while dozens of smORF peptides have been recently described as essential players in biological processes, many are reported to be potential non-functional products of junk DNA under pervasive translation, leading to the question: from what perspective is this lack of function assessed? In this context, it was recently suggested that non-functional smORF peptides might play a major role during *de novo* protein coding gene birth, but the evolutionary mechanism is still unclear. Thus, the role of pervasive translation of smORFs in molecular evolution remains puzzling. Here, we present interesting questions for debate and further investigation about the perspective of non-functional smORF peptides as underappreciated hotspots of molecular evolution in eukaryotes.

## SMALL OPEN READING FRAMES: A SUBTOPIC IN THE DISCUSSION OF JUNK DNA FUNCTION

With respect to the evolution of molecular function, part of the DNA elements accumulate mutations by genetic drift; thus, the evolution of these elements is non-adaptive and neutral (Ohta, 2002). In some cases, the amount of neutrally evolving elements in junk DNA are analogous to the items on a menu available to natural selection (Knibbe et al., 2007; Faulkner and Carninci, 2009; Lynch et al., 2011). Interestingly, it was reported by the ENCODE consortium (the Encyclopedia of DNA Elements) that most of the human junk DNA exhibits some type of biochemical activity (ENCODE Project Consortium, 2012), but lacking adaptive relevance and selective pressure (Doolittle, 2013; Graur et al., 2013). Importantly, junk DNA represents 75–90% of the human genome (Graur, 2017).

Part of the junk DNA menu is composed of neutrally evolving smORF peptides. For instance, thousands of non-coding RNAs are generated by the extensive transcription coverage on junk DNA (ENCODE Project Consortium, 2007). Increasing evidence shows that thousands of smORFs undergo pervasive translation in transcripts annotated as non-coding or in untranslated regions (UTR) of mRNAs (e.g., Aspden et al., 2014; Ingolia et al., 2014). Interestingly, non-coding RNAs and ORFs lacking homologs were reported to be candidates for *de novo* evolution of protein coding genes (Tautz and Domazet-Lošo, 2011). Moreover, it was recently suggested that neutrally evolving smORF peptides might play a major role in this process (Ruiz-Orera et al., 2018), but the

evolutionary mechanism remains to be determined (Ruiz-Orera et al., 2018; Singh and Wurtele, 2020). In this context, two previously proposed concepts used to discuss molecular function evolution are at the core of the junk DNA debate: “causal roles” and “selected effects” (Doolittle and Brunet, 2017), which will be discussed here in the context of smORFs and protein coding gene birth.

The “causal role” describes the activity performed by a neutrally evolving element by chance. For example, a hypothetical genomic sequence generated by a random nucleotide mutation to resemble a TATA box may be recognized and bound by transcription factors but does not trigger gene transcription (Griffiths, 2009; Graur et al., 2013). In other words, “causal roles” are non-adaptive phenotypes, their emergence is random, and they tend to rapidly disappear during evolution. On the other hand, “selected effects” describe the acquisition of adaptive phenotypes based on natural selection (Graur et al., 2013), such as canonical TATA boxes or ORFs that are translated into important proteins. In other words, “selected effects” are functionally relevant for cells.

Importantly, while natural selection drives adaptive evolution (selected effects), it is widely accepted that genetic drift drives junk DNA evolution, as well as the synonymous modifications in coding DNA sequences (CDS) and mutations in UTRs of mRNAs (Ridley, 2004).

## DISCUSSION

Applying the aforementioned evidence and concepts, we discuss here a possible eukaryotic mechanism by which neutrally evolving smORFs advance proteome evolution and the evolutionary significance of smORFs.

Firstly, part of the roles performed by neutrally evolving smORF peptides possibly transit from “causal roles” to “selected effects” under environmental pressure, thereby exposing their neutral phenotypes to natural selection and triggering the evolution of new coding genes. Thus, when neutral smORF peptides are selected, they are no longer neutral (Ruiz-Orera et al., 2018). In other words, neutral smORF peptides may be special entrees on the junk DNA menu that are available for natural selection (Figure 1A).

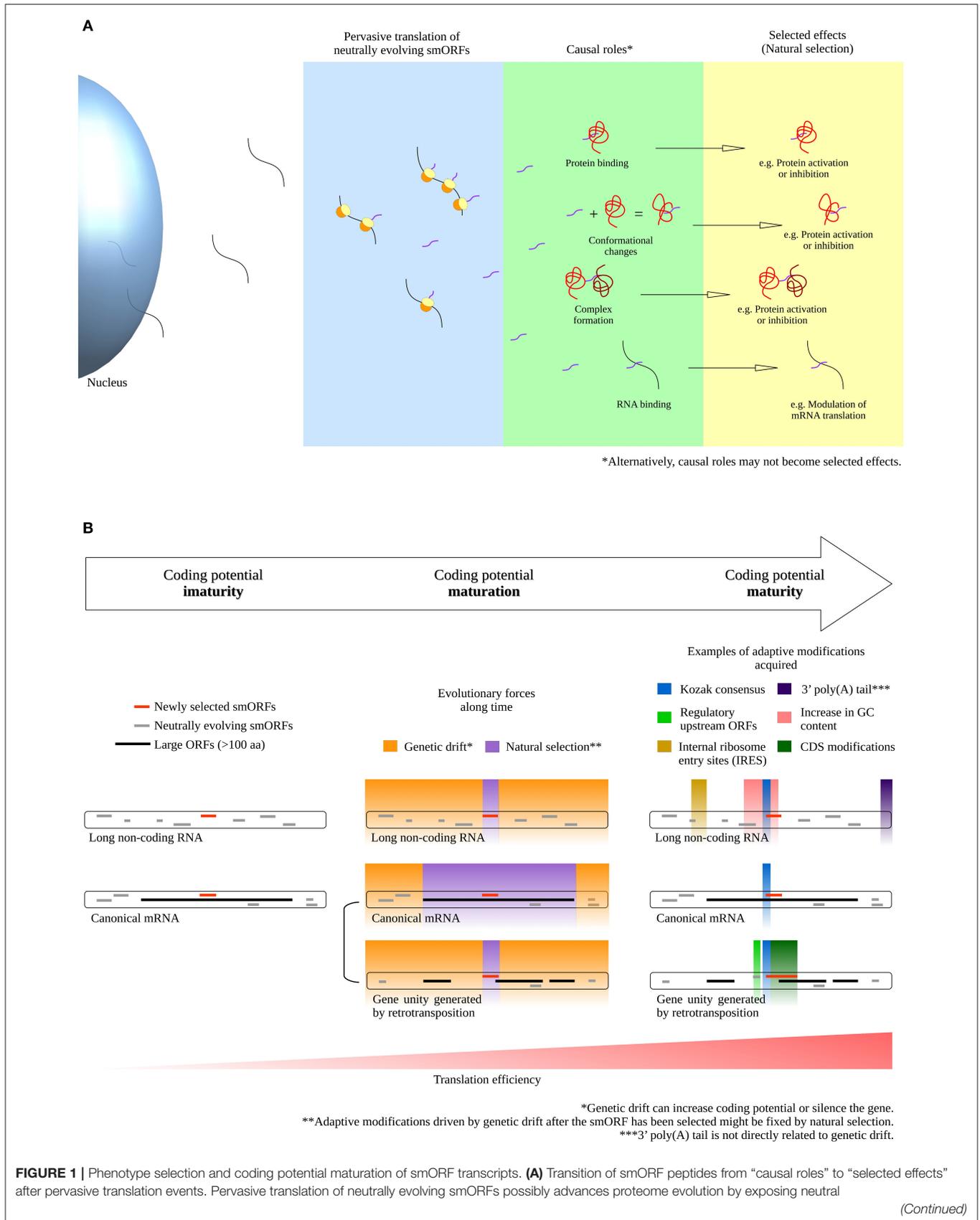
Upon smORFs being selected for, they probably contain low adaptive relevance due to their non-coding transcript characteristics, such as low translation rate, lack of 3'-terminal processing and other suboptimal coding features (non-coding RNA features are reviewed in Quinn and Chang, 2016). This hypothesis is based on the fact that hundreds of smORFs are described as highly conserved but display low expression, low translation efficiency and are observed in transcripts with non-coding characteristics (Cabili et al., 2011; Aspden et al., 2014; Bazzini et al., 2014). However, the nearly neutral theory (Ohta, 2002) suggests that non-coding parts of fixed smORF transcripts are modified by random genetic drift, in some cases, producing small advantageous (or disadvantageous) adaptive effects throughout evolution; thus, we propose that, at a certain point, these modifications refine and elevate the coding potential

of smORF transcripts and consequently enhance the adaptive relevance of their peptides, as seen in a large number of important smORF peptides recently discovered (e.g., Magny et al., 2013; Anderson et al., 2015; Laressergues et al., 2015; Nelson et al., 2016; Pengpeng et al., 2017; Kim et al., 2018; Polycarpou-Schwarz et al., 2018; Chugunova et al., 2019; Tobias-Santos et al., 2019; Pang et al., 2020; Vassallo et al., 2020). Importantly, the acquisition of several optimal coding features might be favored after the smORF has been selected for, because modifications driven by genetic drift could be fixed by natural selection if they improve the translation efficiency of the newly selected smORF. Before the smORF has been selected for, eventual optimal coding features acquired in the nucleotide sequence could rapidly disappear during genetic drift evolution without fixation. Alternatively, nucleotide changes may negatively affect the coding potential and silence a gene. Optimal coding features include structural stabilization, emergence of Kozak consensus, internal ribosome entry sites (IRES), coverage by enhancers and, in some cases, the elongation of coding smORFs to enlarge the CDSs (Figure 1B). Recently, Couso and Patraquim (2017) proposed that at least a portion of functional smORFs are potential *de novo* precursors of large CDSs via a stop codon mutation pattern called “CDS elongation.”

Considering the supposition that the action of evolution is gradual, we propose that the aforementioned process be called “coding potential maturation” (Figure 1B). For example, smORF translation is widely reported in transcripts with long non-coding RNA (lncRNA) characteristics (Crappé et al., 2013; Ingolia et al., 2014; Ji et al., 2015; Mackowiak et al., 2015; Li et al., 2018; Lu et al., 2019). These lncRNAs exhibit smORF conservation in divergent species, hinting at natural selection fixation and indicating coding immaturity.

Another potential pathway of coding gene generation occurs via alternative smORFs in UTRs or overlapping the reference CDS of canonical mRNAs. In this scenario, alternative smORFs undergo pervasive translation or the act of translation itself is important for cis-regulatory purposes (Vanderperre et al., 2013; Wu et al., 2020). If the “causal roles” performed by neutrally evolving smORF peptides become “selected effects,” the alternative smORFs would generate independent gene units by retrotransposition, or they would be fixed as alternative smORFs in the original transcripts (Figure 1B). Hence, during retrotransposition events, at least a portion of the transcripts investigated on the basis of pseudogenization may, in fact, represent the maturation of new coding genes, as suggested by a report that pseudogenes can be translated into highly conserved smORF peptides (Ji et al., 2015).

smORFs might be sequence reservoirs potentially activated during the evolution of new phenotypic variations, especially during speciation. Importantly, speciation events have been associated with the evolution of new molecular phenotypes and new relationships with the environment (Bao et al., 2018). Thus, the amount of junk DNA and lncRNAs in cells deserves investigation not only as a random accumulation of sequences and translational noise but also as a repository of substrates to advance the evolution of new coding genes. Interestingly, polyploidization, or whole genome duplication



**FIGURE 1 |** phenotypes to natural selection under environmental pressure. **(B)** Scheme for coding potential maturation, a hypothetical mechanism that increase the translation efficiency of a mRNA after a smORF has been selected for (selected effect) in a transcript with suboptimal coding features. On the left, coding potential immaturity; in the middle, coding potential maturation; on the right, coding potential maturity. During the coding potential immaturity phase, newly selected smORFs are observed in transcripts with suboptimal coding features, either in long non-coding RNAs or as alternative smORFs in canonical mRNAs. Although canonical mRNAs exhibit optimal coding features, alternative smORFs are usually secondarily or pervasively translated; thus, some alternative smORFs may reside in suboptimal coding regions. During the coding potential maturation phase, natural selection and genetic drift may act in different parts of a transcript. While natural selection acts by fixing the selected parts, genetic drift acts by changing the non-coding parts of a transcript, as postulated by the nearly neutral theory (Ohta, 2002). Natural selection promotes fine-tuned adjustments to the selected phenotypes, such as synonymous mutations and CDS modifications. Genetic drift can establish adaptive mutations in a transcript by evolving sequences that potentially increase smORF translation, such as the Kozak consensus, regulatory upstream ORFs, internal ribosome entry sites (IRES) and increases in GC content. Additionally, other adaptive modifications not directly related to sequence mutations in transcripts might increase smORF expression, such as the 5' cap, 3' poly(A) tail, cis-regulatory elements in the genome and, in the case of alternative smORFs, independent gene unit generation by retrotransposition. Importantly, the acquisition of optimal coding features might be favored after the smORF has been selected for, because modifications driven by genetic drift could be fixed by natural selection if they improve the translation efficiency of the newly selected smORF. Before the smORF has been selected for, eventual optimal coding features acquired could rapidly disappear during genetic drift evolution without fixation. Alternatively, mutations evolved by genetic drift can silence the gene. Finally, smORFs reach the coding potential maturity phase when optimal coding features are acquired and translation efficiency increases. Consequently, the translation rate of smORF peptides is largely increased upon completion of the described process, contributing to the establishment of molecular innovations and protein coding gene birth.

(WGD) events, have been correlated with an increase in the adaptive potential of cells and organisms exposed to stressful conditions (Van De Peer et al., 2017). Unfortunately, thus far, studies of WGD have neglected the role and retention of smORFs during evolution, probably due to methodological difficulties in smORF identification.

However, the sequencing of several genomes based on comparative approaches has recently opened new avenues for smORF research. For instance, recent evolutionary studies performed by our group on the smORFs in the *mille-pattes/tarsalles/polished rice (mlpt)* gene, the most well-known smORF-containing gene in insects (Savard et al., 2006; Kondo et al., 2007; Pueyo and Couso, 2008, 2011; Cao et al., 2017; Ray et al., 2019), showed that a new ~80 amino acid smORF (smHemiptera) appeared during Hemiptera evolution (Tobias-Santos et al., 2019). Thus, this smORF in the polycistronic *mlpt* mRNA has been conserved for over 250 million years in the group, and it is not present in the genomes of other insect orders. We expect that new comparative analyses of genomes in the future will yield additional examples of order-specific smORFs, which might constitute an underappreciated reservoir of new genes and evolutionary innovations.

In summary, the study of smORFs has been considerably increasing during the last 5 years because of recent discoveries of important smORF peptides. Accordingly, the advent of ribosome profiling has allowed the discovery of many neutrally evolving and potentially non-functional smORFs undergoing pervasive translation, whose significance remains

to be determined (Crappé et al., 2013; Aspden et al., 2014; Bazzini et al., 2014; Olexiuk et al., 2016). In this context, the intriguing question is posed: why would cells spend energy on transcription and translation of neutral and non-functional elements? There is probably more than one answer; however, considering the subjects discussed in this paper, we propose the following perspective: what if the pervasive translation of neutrally evolving smORF peptides composes an elegant mechanism to advance proteome evolution, especially during speciation events? If it does, then non-functional smORF peptides display an important function in an evolutionary sense. Based on this discussion, we suggest that the concept of functionality be revised in the context of smORFs.

## AUTHOR CONTRIBUTIONS

DG-A and RN contributed equally to the writing of this manuscript. RN contributed to funding acquisition. All authors contributed to the article and approved the submitted version.

## FUNDING

RN was supported by CNPq (307952/2017-7 and 431354/2016-2) and FAPERJ (E-26/210-150/2016, E-26/203.298/2016, E-26/202.605/2019, and E-26/211.169/2019). DG-A was a master's student of PPG-PRODBIO-UFRJ/Macaé (CAPES scholarship).

## REFERENCES

- Anderson, D. M., Anderson, K. M., Chang, C. L., Makarewich, C. A., Nelson, B. R., McAnally, J. R., et al. (2015). A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell* 160, 595–606. doi: 10.1016/j.cell.2015.01.009
- Aspden, J. L., Eyre-Walker, Y. C., Phillips, R. J., Amin, U., Mumtaz, M. A. S., Brocard, M., et al. (2014). Extensive translation of small open reading frames revealed by poly-ribo-seq. *eLife* 3:e03528. doi: 10.7554/eLife.03528
- Bao, R., Dia, S. E., Issa, H. A., Alhusein, D., and Friedrich, M. (2018). Comparative evidence of an exceptional impact of gene duplication on the developmental evolution of *Drosophila* and the higher Diptera. *Front. Ecol. Evol.* 6:63. doi: 10.3389/fevo.2018.00063
- Bazzini, A. A., Johnstone, T. G., Christiano, R., Mackowiak, S. D., Obermayer, B., Fleming, E. S., et al. (2014). Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *EMBO J.* 33, 981–993. doi: 10.1002/emboj.201488411
- Cabili, M., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., et al. (2011). Integrative annotation of human large intergenic noncoding RNAs

- reveals global properties and specific subclasses. *Genes Dev.* 25, 1915–1927. doi: 10.1101/gad.17446611
- Cao, G., Gong, Y., Hu, X., Zhu, M., Liang, Z. C., Huang, L., et al. (2017). Identification of tarsal-less peptides from the silkworm *Bombyx mori*. *Appl. Microbiol. Biotechnol.* 102, 1809–1822. doi: 10.1007/s00253-017-8708-4
- Chugunova, A., Loseva, E., Mazin, P., Mitina, A., Navalayeu, T., Bilan, D., et al. (2019). LINC00116 codes for a mitochondrial peptide linking respiration and lipid metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 116, 4940–4945. doi: 10.1073/pnas.1809105116
- Couso, J., and Patraquim, P. (2017). Classification and function of small open reading frames. *Nat. Rev. Mol. Cell Biol.* 18, 575–589. doi: 10.1038/nrm.2017.58
- Crappé, J., Crieckinge, W. V., Trooskens, G., Hayakawa, E., Luyten, W., Baggerman, G., et al. (2013). Combining *in silico* prediction and ribosome profiling in a genome-wide search for novel putatively coding sORFs. *BMC Genomics* 14:648. doi: 10.1186/1471-2164-14-648
- Doolittle, W. F. (2013). Is junk DNA bunk? A critique of ENCODE. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5294–5300. doi: 10.1073/pnas.1221376110
- Doolittle, W. F., and Brunet, T. D. P. (2017). On causal roles and selected effects: our genome is mostly junk. *BMC Biol.* 15:116. doi: 10.1186/s12915-017-0460-9
- ENCODE Project Consortium (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447, 799–816. doi: 10.1038/nature05874
- ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. doi: 10.1038/nature11247
- Faulkner, G. J., and Carninci, P. (2009). Altruistic functions for selfish DNA. *Cell Cycle* 8, 2895–900. doi: 10.4161/cc.8.18.9536
- Graur, D. (2017). An upper limit on the functional fraction of the human genome. *Genome Biol. Evol.* 9, 1880–1885. doi: 10.1093/gbe/evx121
- Graur, D., Zheng, Y., Price, N., Azevedo, R. B. R., Zufall, R. A., Elhaik, E., et al. (2013). On the immortality of television sets: “Function” in the human genome according to the evolution-free gospel of ENCODE. *Genome Biol. Evol.* 5, 578–590. doi: 10.1093/gbe/evt028
- Griffiths, P. E. (2009). In what sense does “nothing make sense except in the light of evolution?” *Acta Biotheor.* 57:11. doi: 10.1007/s10441-008-9054-9
- Ingolia, N. T., Brar, G. A., Stern-Ginossar, N., Harris, M. S., Talhouarne, G. J. S., Jackson, S. E., et al. (2014). Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell Rep.* 8, 1365–1379. doi: 10.1016/j.celrep.2014.07.045
- Ji, Z., Song, R., Regev, A., and Struhl, K. (2015). Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *eLife* 4:e08890. doi: 10.7554/eLife.08890
- Kim, K. H., Son, J. M., Benayoun, B. A., and Lee, C. (2018). The mitochondrial-encoded peptide MOTS-c translocates to the nucleus to regulate nuclear gene expression in response to metabolic stress. *Cell Metab.* 28, 516.e7–524.e7. doi: 10.1016/j.cmet.2018.06.008
- Knibbe, C., Coulon, A., Mazet, O., Fayard, J. M., and Beslon, G. (2007). A long-term evolutionary pressure on the amount of noncoding DNA. *Mol. Biol. Evol.* 24, 2344–2353. doi: 10.1093/molbev/msm165
- Kondo, T., Hashimoto, Y., Kato, K., Inagaki, S., Hayashi, S., and Kageyama, Y. (2007). Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. *Nat. Cell Biol.* 9, 660–665. doi: 10.1038/ncb1595
- Lauressergues, D., Couzigou, J. M., Clemente, H. S., Martinez, Y., Dunand, C., Bécard, G., et al. (2015). Primary transcripts of microRNAs encode regulatory peptides. *Nature* 520, 90–93. doi: 10.1038/nature14346
- Li, H., Xiao, L., Zhang, L., Wu, J., Wei, B., Sun, N., et al. (2018). FSPP: A tool for genome-wide prediction of smORF-encoded peptides and their functions. *Front. Genet.* 9:96. doi: 10.3389/fgene.2018.00096
- Lu, S., Zhang, J., Lian, X., Sun, L., Meng, K., Chen, Y., et al. (2019). A hidden human proteome encoded by ‘non-coding’ genes. *Nucleic Acids Res.* 47, 8111–8125. doi: 10.1093/nar/gkz646
- Lynch, M., Bobay, L. M., Catania, F., Gout, J. F., and Rho, M. (2011). The repatterning of eukaryotic genomes by random genetic drift. *Annu. Rev. Genomics Hum. Genet.* 12, 347–366. doi: 10.1146/annurev-genom-082410-101412
- Mackowiak, S. D., Zaubner, H., Bielow, C., Thiel, D., Kutz, K., Calviello, L., et al. (2015). Extensive identification and analysis of conserved small ORFs in animals. *Genome Biol.* 16:179. doi: 10.1186/s13059-015-0742-x
- Magny, E. G., Pueyo, J. I., Pearl, F. M. G., Cespedes, M. A., Niven, J. E., Bishop, S. A., et al. (2013). Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. *Science* 341, 1116–1120. doi: 10.1126/science.1238802
- Nelson, B. R., Makarewich, C. A., Anderson, D. M., Winders, B. R., Troupes, C. D., Wu, F., et al. (2016). A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science* 351, 271–275. doi: 10.1126/science.aad4076
- Ohta, T. (2002). Near-neutrality in evolution of genes and gene regulation. *Proc. Natl. Acad. Sci. U.S.A.* 99, 16134–16137. doi: 10.1073/pnas.252626899
- Olexiouk, V., Crappé, J., Verbruggen, S., Verhegen, K., Martens, L., and Menschaert, G. (2016). sORFs.org: a repository of small ORFs identified by ribosome profiling. *Nucleic Acids Res.* 44, D324–D329. doi: 10.1093/nar/gkv1175
- Pang, Y., Liu, Z., Han, H., Wang, B., Li, W., Mao, C., et al. (2020). Peptide SMIM30 promotes HCC development by inducing SRC/YES1 membrane anchoring and MAPK pathway activation. *J. Hepatol.* doi: 10.1016/j.jhep.2020.05.028. [Epub ahead of print].
- Pengpeng, B., Ramirez-Martinez, A., Li, H., Cannavino, J., McAnally, J. R., Shelton, J. M., et al. (2017). Control of muscle formation by the fusogenic micropeptide myomixer. *Science* 356, 323–327. doi: 10.1126/science.aam9361
- Polycarpou-Schwarz, M., Groß, M., Mestdagh, P., Schott, J., Grund, S. E., Hildenbrand, C., et al. (2018). The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation. *Oncogene* 37, 4750–4768. doi: 10.1038/s41388-018-0281-5
- Pueyo, J. I., and Couso, J. P. (2008). The 11-aminoacid long tarsal-less peptides trigger a cell signal in *Drosophila* leg development. *Dev. Biol.* 324, 192–201. doi: 10.1016/j.ydbio.2008.08.025
- Pueyo, J. I., and Couso, J. P. (2011). Tarsal-less peptides control Notch signalling through the shavenbaby transcription factor. *Dev. Biol.* 355, 183–1936. doi: 10.1016/j.ydbio.2011.03.033
- Quinn, J. J., and Chang, H. Y. (2016). Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* 17, 47–62. doi: 10.1038/nrg.2015.10
- Ray, S., Rosenberg, M. I., Chanut-Delalande, H., Decaras, A., Schwertner, B., Toubiana, W., et al. (2019). The mlpt/Ubr3/Svb module comprises an ancient developmental switch for embryonic patterning. *eLife* 8:e39748. doi: 10.7554/eLife.39748
- Ridley, M. (2004). *Evolution, 3rd edn.* Oxford: Blackwell Pub.
- Ruiz-Orera, J., Verdaguier-Grau, P., Villanueva-Cañas, J. L., Messeguer, X., and Albà, M. M. (2018). Translation of neutrally evolving peptides provides a basis for *de novo* gene evolution. *Nat. Ecol. Evol.* 2, 890–896. doi: 10.1038/s41559-018-0506-6
- Savard, J., Marques-Souza, H., Aranda, M., and Tautz, D. (2006). A segmentation gene in *Tribolium* produces a polycistronic mRNA that codes for multiple conserved peptides. *Cell* 126, 559–569. doi: 10.1016/j.cell.2006.05.053
- Singh, U., and Wurtele, E. S. (2020). How new genes are born. *eLife* 2020:e55136. doi: 10.7554/eLife.55136
- Tautz, D., and Domazet-Lošo, T. (2011). The evolutionary origin of orphan genes. *Nat. Rev. Genet.* 12, 692–702. doi: 10.1038/nrg3053
- Tobias-Santos, V., Guerra-Almeida, D., Murry, F., Ribeiro, L., Berni, M., Araujo, H., et al. (2019). Multiple roles of the polycistronic gene *Tarsal-Less/Mille-Pattes/Polished-Rice* during embryogenesis of the kissing bug *Rhodnius prolixus*. *Front. Ecol. Evol.* 7:379. doi: 10.3389/fevo.2019.00379
- Van De Peer, Y., Mizrachi, E., and Marchal, K. (2017). The evolutionary significance of polyploidy. *Nat. Rev. Genet.* 18, 411–424. doi: 10.1038/nrg.2017.26
- Vanderperre, B., Lucier, J. F., Bissonnette, C., Motard, J., Tremblay, G., Vanderperre, S., et al. (2013). Direct detection of alternative open reading

- frames translation products in human significantly expands the proteome. *PLoS ONE* 8:e70698. doi: 10.1371/journal.pone.0070698
- Vassallo, A., Palazzotto, E., Renzone, G., Botta, L., Faddetta, T., Scaloni, A., et al. (2020). The *Streptomyces coelicolor* small ORF *trpM* stimulates growth and morphological development and exerts opposite effects on actinorhodin and calcium-dependent antibiotic production. *Front. Microbiol.* 11:224. doi: 10.3389/fmicb.2020.00224
- Wu, Q., Wright, M., Gogol, M. M., Bradford, W. D., Zhang, N., and Bazzini, A. (2020). Translation of small downstream ORFs enhances translation of canonical main open reading frames. *EMBO J.* 39:e104763. doi: 10.15252/embj.2020104763

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Guerra-Almeida and Nunes-da-Fonseca. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.