



Unveiling the Impact of the Genomic Architecture on the Evolution of Vertebrate microRNAs

Gustavo S. França^{1*}, Ludwig C. Hinske², Pedro A. F. Galante³ and Maria D. Vibranovski^{1*}

¹ Departamento de Genética e Biologia Evolutiva, Universidade de São Paulo, São Paulo, Brazil, ² Department of Anesthesiology, Clinic of the University of Munich, Ludwig Maximilian University of Munich, Munich, Germany, ³ Centro de Oncologia Molecular, Hospital Sírio-Libanês, São Paulo, Brazil

OPEN ACCESS

Edited by:

Scott V. Edwards,
Harvard University, USA

Reviewed by:

Antonio Marco,
University of Essex, UK
Bastian Fromm,
Oslo University Hospital, Norway
Danillo Pinhal,
São Paulo State University (UNESP),
Brazil

*Correspondence:

Gustavo S. França
gustavo.starvaggi@franca@nyumc.org
Maria D. Vibranovski
mdv@ib.usp.br

† Present address:

Gustavo S. França,
Institute for Computational Medicine,
NYU School of Medicine, 430 East
29th Street, New York, NY, USA

Specialty section:

This article was submitted to
Evolutionary and Population Genetics,
a section of the journal
Frontiers in Genetics

Received: 09 December 2016

Accepted: 09 March 2017

Published: 21 March 2017

Citation:

França GS, Hinske LC, Galante PAF
and Vibranovski MD (2017) Unveiling
the Impact of the Genomic
Architecture on the Evolution
of Vertebrate microRNAs.
Front. Genet. 8:34.
doi: 10.3389/fgene.2017.00034

Eukaryotic genomes frequently exhibit interdependency between transcriptional units, as evidenced by regions of high gene density. It is well recognized that vertebrate microRNAs (miRNAs) are usually embedded in those regions. Recent work has shown that the genomic context is of utmost importance to determine miRNA expression in time and space, thus affecting their evolutionary fates over long and short terms. Consequently, understanding the inter- and intraspecific changes on miRNA genomic architecture may bring novel insights on the basic cellular processes regulated by miRNAs, as well as phenotypic evolution and disease-related mechanisms.

Keywords: intragenic, intergenic region, new and old miRNAs, host gene, target interactions, expression breadth

INTRODUCTION

Recent genome-wide projects have revealed an outstanding transcriptome diversity, especially of non-coding RNAs (ncRNAs), as well as a wealth of regulatory mechanisms and gene product interactions that compound the molecular basis of phenotypes (Carninci et al., 2005; Mele et al., 2015). A notable feature that soon became clear is the interleaved nature of eukaryotic genomes, despite their typical large sizes. This means that a particular genomic region can be suited for different purposes, with an extensive overlap of transcriptional units either in sense or antisense DNA strands (Kapranov et al., 2007).

The interleaved model opens up numerous possibilities for regulatory mechanisms. For instance, products of antisense transcription, which is believed to occur in more than 30% of gene loci in humans (Galante et al., 2007), can regulate gene activity through many different ways (reviewed in Pelechano and Steinmetz, 2013). In the interleaved genome, transcription units may show high interdependency, whereby neighboring or overlapping genes can be co-regulated by shared regulatory elements; yet, structural changes in the chromatin environment can also influence their expression coordinately (Mellor et al., 2016). Complex transcriptional networks thus emerge from a modular architecture that can either be shaped by evolutionary advantages and constraints (Mercer and Mattick, 2013), but also as a result of neutral processes (Graur et al., 2015). Such interleaved architecture is particularly striking in regard to microRNAs (miRNAs). Ever since the first large-scale studies on their genomic organization (Rodriguez et al., 2004), it is commonly observed that these small non-coding RNAs overlap to protein-coding genes, with vertebrate miRNAs mapping to intronic regions more than expected by chance (Baskerville and Bartel, 2005; Hinske et al., 2010, 2014; Campo-Paysaa et al., 2011; Meunier et al., 2013). As they comprise an essential class of gene expression regulators in basic biological processes and diseases, genomic

context analyses are pivotal to uncover unique aspects of miRNA biology. Here, we discuss recent advances in this topic focusing on the importance of the genomic context to miRNA expression and their target interactions. In this framework, we highlight the evolutionary consequences for the fixation of newly emerged miRNAs and functional properties arising from miRNA–genomic context relationships over long- and short-evolutionary terms.

THE IMPACT OF THE GENOMIC CONTEXT ON miRNA EXPRESSION AND FUNCTION

As any other gene, the evolutionary processes that gives rise to new miRNAs – mainly by duplication or *de novo* origin (Berezikov, 2011; Meunier et al., 2013) – takes place on certain regions of the genome that may overlap or not to preexisting gene loci. In a recent study, Meunier et al. (2013) showed that all vertebrate species analyzed (Chicken, Platypus, Opossum, Mouse, Macaque, and Human) have a significant excess of intragenic miRNAs, with on average 54% of them overlapping to introns. Curiously, the proportions of intronic miRNAs are even higher for those of recent origin, suggesting that introns are hotspots for new miRNA origination. Moreover, the transcriptional orientation of intragenic miRNAs is highly biased (~80%) toward the same strand orientation of their host genes (Rodriguez et al., 2004; Campo-Paysaa et al., 2011; Meunier et al., 2013; Hinske et al., 2014).

Given the large size of vertebrate genomes, why do miRNAs apparently have such preference to emerge in intragenic regions? Which evidences support the role of natural selection shaping this pattern, and what advantages miRNAs might take from such genomic organization? To address these questions, França et al. (2016) investigated the patterns of emergence and expression of human miRNAs along the vertebrate evolution considering the evolutionary origin of their host genes, i.e., whether miRNAs are intergenic, mapped to old protein-coding genes (originated before fish and tetrapods divergence), or to young protein-coding genes (originated after the divergence). Similar to previous studies (Iwama et al., 2013), it was shown that most human miRNAs (~70%) have a relatively recent origin, emerging in the primate order. Though an interesting pattern was revealed, the majority of those young miRNAs are intragenic and preferentially embedded within old host genes, even when controlled by host gene length (including intronic region) and expression level. Expression breadth analyses showed that young miRNAs hosted by old genes were more broadly expressed (expression in more tissues) than their intergenic counterparts. On the other hand, miRNAs hosted by young genes showed a bias to tissue-specific expression when compared to the intergenic ones or those within old genes. The same conclusions held when a very stringent miRNA annotation provided by Fromm et al. (2015) was considered, since several miRBase entries do not represent bonafide miRNAs (Chiang et al., 2010; Taylor et al., 2014; Fromm et al., 2015). It is well established that expression breadth is negatively correlated

with evolutionary rates (Wolf et al., 2009; Park and Choi, 2010), meaning that overall conserved genes are highly and broadly expressed, whereas less conserved genes tend to have low and narrow expression. What turns out is that the expression of intragenic miRNAs is tightly coupled to their genomic environment, especially in regard to the evolutionary ages of their host genes. In a mechanistic way, this is clearly connected with the co-expression of miRNA–host gene pairs by shared regulatory elements, a very well-documented event (Baskerville and Bartel, 2005; Ozsolak et al., 2008; Marsico et al., 2013). Hence, the maintenance of miRNAs embedded in genic regions may be indicative of some evolutionary constraint, since young and older intragenic miRNAs are biased toward host gene sense orientation, as well as preferential emergence within old host genes. In addition, same age miRNAs show differential expression breadth depending on their genomic context, a pattern that is maintained not only during recent (e.g., primates) but also over longer periods. Such pattern is observed for miRNAs originated in amniotes (e.g., chicken) or in placental mammals (e.g., mouse) presenting higher or lower expression breadth depending on the age of their host genes (França et al., 2016).

In particular for young intragenic miRNAs, being hosted by old genes could be beneficial at least during an initial adaptive phase, because of the expression broadness achieved through a presumably favorable transcriptional environment. Instead of readily relying on the settlement of their own regulatory apparatus, young miRNAs would initially be benefited by their hosts' regulatory elements, albeit they may acquire independent regulation afterward (França et al., 2016). Supporting this notion, it has been suggested that young and middle-aged intragenic miRNAs are more likely to be regulated by shared promoters, whereas old miRNAs are frequently regulated by their independent intronic promoters (Marsico et al., 2013). In addition, as old host genes provide higher expression breadth for those young miRNAs, it would, in principle, increase the opportunities for new target interactions in different tissues. From such perspective, the host transcriptional environment could facilitate the initial expression of young miRNAs and thereafter contribute to the process of miRNA functionalization.

The location of a gene in the genome is clearly related to its expression, as revealed by transgene insertion experiments (Mlynárová et al., 2002) and global expression analyses of gene neighborhoods (Caron et al., 2001; Purmann et al., 2007; Michalak, 2008). Nevertheless, some of the observed expression changes in gene vicinity may not be subjected to selection, but rather it would be a consequence of expression changes in a close gene under strong selection. Recently, Ghanbarian and Hurst (2015) demonstrated that expression changes in humans, relative to the human–chimp common ancestor, coordinately drive changes in expression of the neighbors of a focal gene, and that this effect is stronger as the distance between genes are shorter (<100kbp). Therefore, the genomic context still may yield important effects on the expression, and perhaps the fixation of novel miRNAs that are not under direct selection.

EVOLUTIONARY CONSERVATION AND NOVELTIES FROM miRNAs' GENOMIC CONTEXT

The phylogenetic distribution of miRNAs in vertebrates is distinguished by the presence of deeply conserved and abundant clade or species-specific repertoires (Berezikov et al., 2006; Wheeler et al., 2009; Meunier et al., 2013; Fromm et al., 2015). Although the evolution of miRNA sequences have been investigated (Lyu et al., 2014; Ninova et al., 2014), the conserved patterns and evolutionary innovations that arose due to interspecific differences in the genomic context are largely underexplored. One of the few studies to address this issue compared the genomic location and expression of ~100 miRNAs during developmental stages of medaka fish, zebrafish, chicken, and mouse (Ason et al., 2006). It was demonstrated that spatial expression differences can be related to changes either in the miRNA location and copy number variation rather than to sequence divergence (Ason et al., 2006). Actually, the miRNA genomic location is thought to influence their expression divergence, as old- and middle-aged intragenic miRNAs tend to be more similarly expressed among species than intergenic ones (França et al., 2016).

Such kind of expression constraint linked to a conserved genomic context is clearly observed for *miR-490* and its host gene *CHRM2* (França et al., 2016). Homologous sequences of *miR-490* are found across amniotes, with identical mature sequences from human to chicken. Gene order and location of *miR-490* in the second intron of *CHRM2* are also preserved (Figure 1A). Although *miR-490* is annotated as intergenic in chicken, predicted transcripts with an intron overlapping *miR-490* are annotated. Expression analyses reveal a strongly conserved pattern among human, rhesus macaque, mouse, and chicken; indicating concomitant expression of *miR-490* and *CHRM2* (Shen et al., 2015) with highest abundance in heart (Figure 1A). The host gene is a muscarinic cholinergic receptor involved in acetylcholine-mediated cardiac chronotropic (heart rate) and inotropic (strength of muscle contraction) effects (Brodde and Michel, 1999), and it has been associated with cardiomyopathy (Zhang et al., 2008). Notably, dysregulation of *miR-490* is also reported in cardiac disease (Cooley et al., 2012) and is involved with proliferation of human coronary artery smooth cells (Sun et al., 2013), suggesting an important functional connection between *miR-490* and *CHRM2*.

As mentioned earlier, the transcriptional environment of host genes may act as a key factor to promote the expression of newly emerged miRNAs. This phenomenon is well illustrated by the primate-specific *miR-625* encoded within *FUT8* (Figure 1B). This host gene is a fucosyltransferase well-conserved throughout animals (Costache et al., 1997; Juliant et al., 2014) that catalyzes fucosylation of glycoproteins, which is essential for activating growth factor receptors (Liu et al., 2011), while its deletion has lethal effects in mice (Wang et al., 2005). *FUT8* is ubiquitously expressed in human tissues (Mele et al., 2015) and *miR-625* seems to follow its host expression pattern (Figure 1B). Considering the young evolutionary age of *miR-625*, its expression levels and breadth are unusually high, thus being frequently altered in

different types of cancer (Zhou et al., 2014; Zheng et al., 2015). It is interesting that *miR-625* has emerged as a promising predictive biomarker in colorectal cancer (Verma et al., 2015; Rasmussen et al., 2016), exhibiting strong association with oxaliplatin (a chemotherapeutic agent used in the treatment of metastatic colorectal cancer) resistance (Rasmussen et al., 2016).

Another singular feature of miRNAs is their frequent occurrence in clusters, originated through tandem or non-local duplications or by *de novo* mutations either in introns or intergenic regions (Berezikov, 2011). Such genomic organization is prone to greatly affect the evolution of newly emerged miRNAs. According to Wang et al. (2016), members of the same cluster tend to exhibit coordinated expression and to target overlapping sets of genes. The authors proposed that clustering arrangement and by developing functions related to the pre-existing miRNAs in the same cluster would help the initial survival of these young miRNAs, until the cluster is settled up by purifying selection. Otherwise, the most usual fate of *de novo* newly emerged miRNAs would be to undergo rapid degeneration. In further support of this “functional co-adaptation” model, clustered young miRNAs indeed present significant signs of adaptive changes that probably drive them to functional constraints associated with the older members of the cluster (Wang et al., 2016).

miRNA-TARGET INTERACTIONS: FUNCTIONAL AND EVOLUTIONARY IMPLICATIONS

If a recently emerged miRNA is expressed and integrated into regulatory networks through consistent and biologically relevant target interactions, it will have more chances to become functional and be retained afterward over long periods (Chen and Rajewsky, 2007; Lyu et al., 2014). Therefore, young miRNAs originated in a genomic context able to boost their expression in multiple tissues would favor target recognition. This idea is consistent with the previous observation that young miRNAs emerged within old host genes are expressed in more tissues and tend to have more predicted targets compared to young intergenic ones (França et al., 2016). We, therefore, suggested a miRNA evolution model that takes into account not only the miRNAs themselves, but also their genomic context (França et al., 2016) (Figure 2). Hence, young miRNAs (or “proto” miRNAs) hosted by old genes would gain higher expression breadth benefited by their host's transcriptional activity, thus enabling many target interactions that, at first glance, are mostly neutral (Chen and Rajewsky, 2007; Nozawa et al., 2016), but could be stabilized by natural selection over time. On the other hand, as young intergenic miRNAs tend to have narrower expression, and apparently less targets to interact with, they could undergo faster degeneration (Figure 2). This degeneration scenario is also most likely to happen with miRNAs emerged within young hosts, because of their general tissue-specific expression signature (França et al., 2016).

Evolutionary sequence conservation has been successfully introduced to reduce the number of false-positive and to increase the signal-to-noise ratio in target predictions. Instead

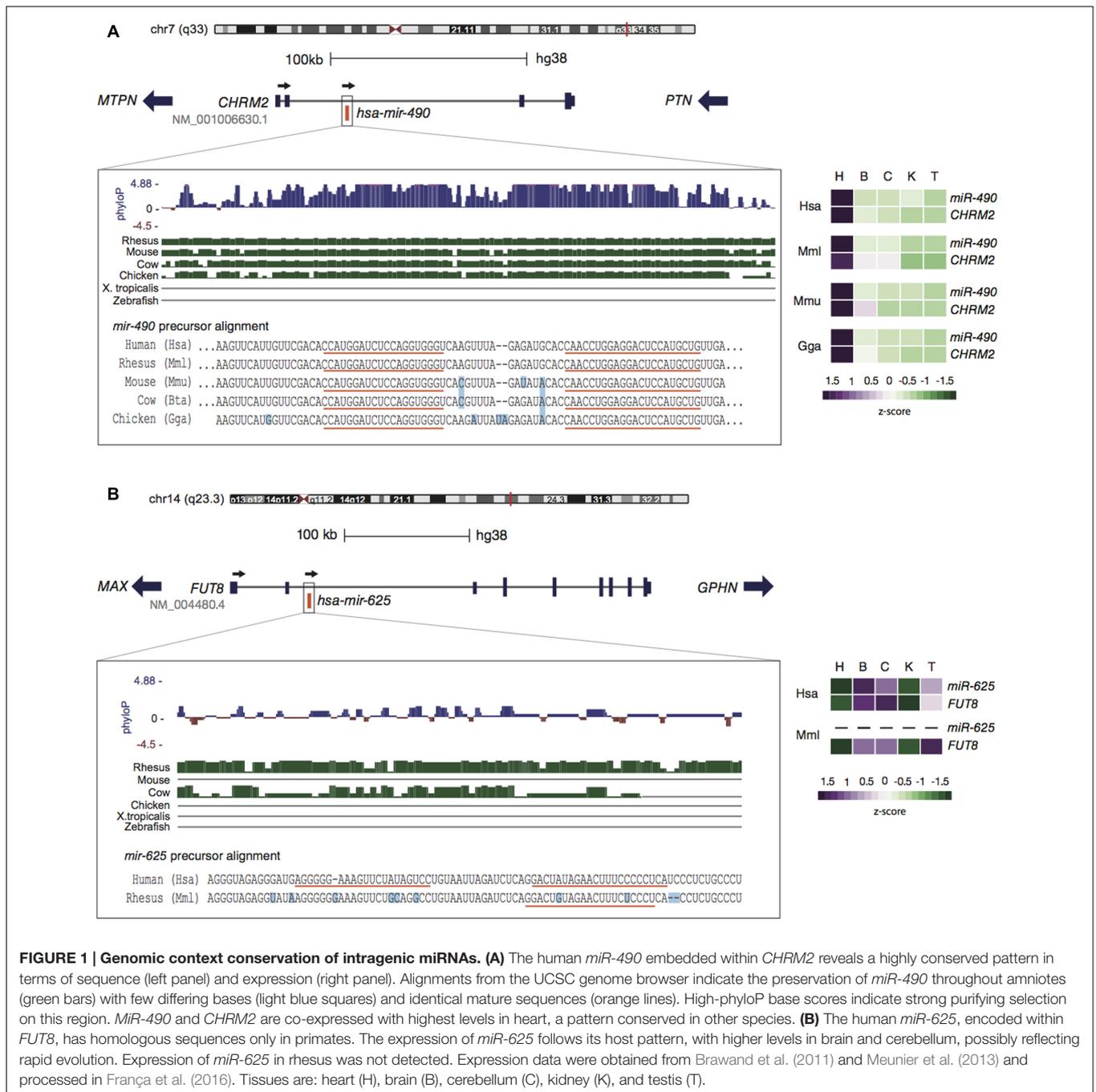


FIGURE 1 | Genomic context conservation of intragenic miRNAs. (A) The human *miR-490* embedded within *CHRM2* reveals a highly conserved pattern in terms of sequence (left panel) and expression (right panel). Alignments from the UCSC genome browser indicate the preservation of *miR-490* throughout amniotes (green bars) with few differing bases (light blue squares) and identical mature sequences (orange lines). High-phyloP base scores indicate strong purifying selection on this region. *miR-490* and *CHRM2* are co-expressed with highest levels in heart, a pattern conserved in other species. **(B)** The human *miR-625*, encoded within *FUT8*, has homologous sequences only in primates. The expression of *miR-625* follows its host pattern, with higher levels in brain and cerebellum, possibly reflecting rapid evolution. Expression of *miR-625* in rhesus was not detected. Expression data were obtained from Brawand et al. (2011) and Meunier et al. (2013) and processed in França et al. (2016). Tissues are: heart (H), brain (B), cerebellum (C), kidney (K), and testis (T).

of helping identifying conserved pathways and relationships among miRNAs and their targets (Hausser and Zavolan, 2014), this requirement comes with a drawback, since it can only be applied to miRNAs and target genes that have conservation data available and which are not species-specific. Indeed, a recent study demonstrated that target sites identified by cross-linking immunoprecipitation data are rarely conserved between distantly related species, but extensive conservation is observed between closely related ones (Xu et al., 2013). Even when considering species-specific sites, there is evidence of selective constraints

compared to non-target sites across the 3'UTR region, suggesting that most of non-conserved targets might be functional at least for a short evolutionary period. A striking example of this condition is the human-specific target site for *miR-183* in the 3vUTR of the transcription factor *FOXO1*, whose regulation altered FOXO1-dependent phenotypes, such as proliferation and migration, in a species-specific manner (McLoughlin et al., 2014). Despite of the recent advances on the characterization of operating mechanisms that guide miRNA–target interactions, we are only on the verge of understanding how newly emerged

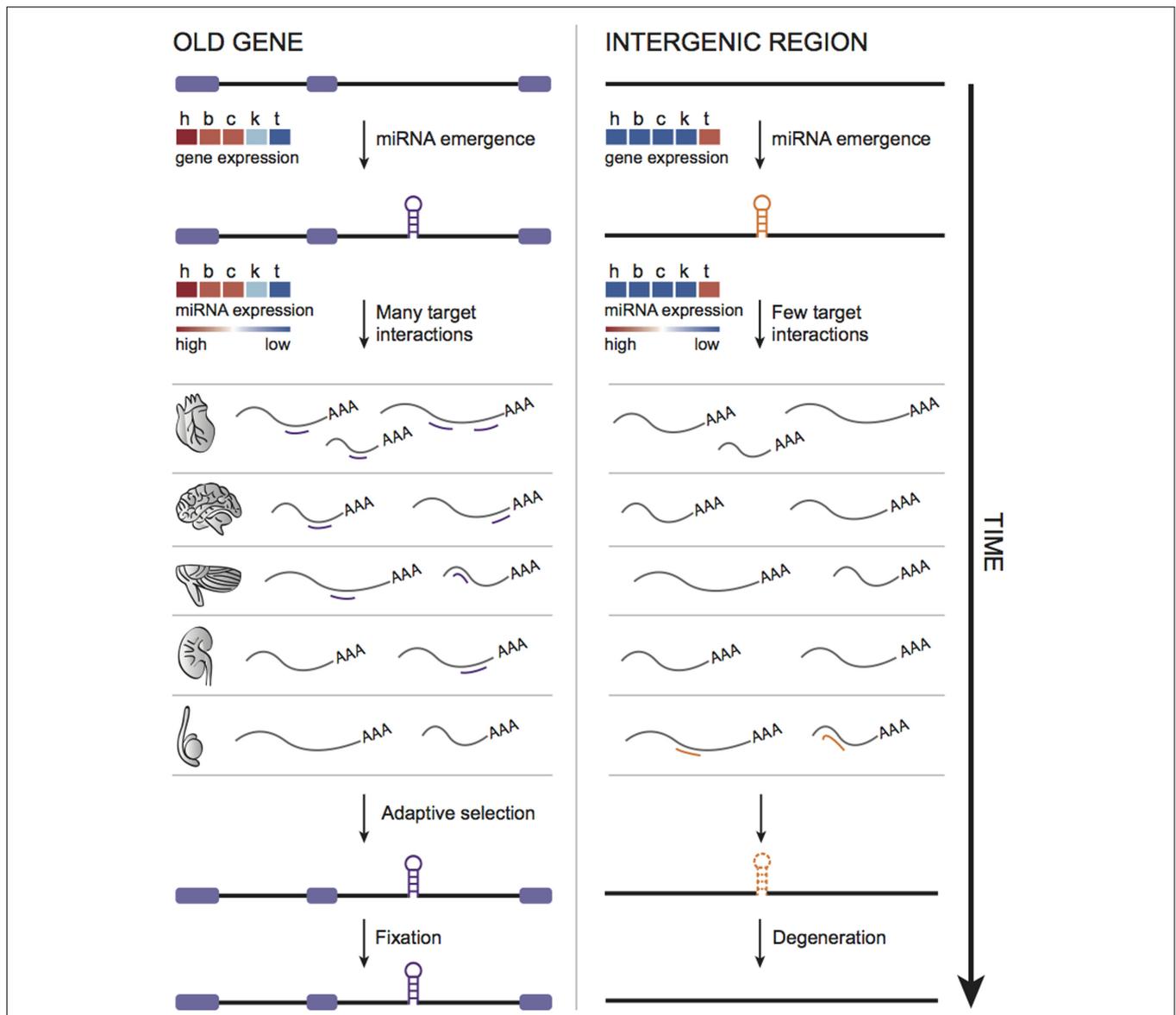


FIGURE 2 | Model of miRNA evolution. Young miRNAs emerged within old genes are expressed in more tissues and, therefore, could interact with diverse set of targets, possibly enhancing the chances of functionalization and fixation through time. In contrast, as young intergenic miRNAs tend to be tissue-specific (likely expressed in testis), very limited target interactions could contribute to their faster degeneration.

miRNAs in different genomic contexts are integrated into regulatory networks, as well as how their novel target interactions contribute to phenotypic plasticity.

POPULATION BIOLOGY PERSPECTIVE FOR THE GENOME ARCHITECTURE OF miRNAS

Population biology studies at the genome level have been proved to be promising tools, enhancing our understanding on how genetic elements are interconnected spatially and temporally (Barrón et al., 2014; Sudmant et al., 2015). Most of

miRNA population studies have focused on the impact of single nucleotide variants localized inside the seed and the mature regions to analyze conservation patterns, target diversification, and differential disease susceptibility (e.g., Barbash et al., 2014; Rawlings-Goss et al., 2014; Gallego et al., 2016). Except for few studies of miRNA expression quantitative trait loci (e.g., Huan et al., 2015), the evolution of miRNA genomic architecture has not been deeply investigated using a population biology framework.

It is still unknown if variation in miRNAs sequence, expression, and target sites across populations are more relevant for uncovering the mechanisms of phenotypic evolution and disease than other genetic variation. On one hand, due to its

folded structure and small size, miRNAs are more likely to emerge *de novo* than novel protein coding genes (Berezikov, 2011). Diversification of miRNA target repertoire may be more prone to appear as result of simple sequence modifications such as direct mutation, seed or hairpin shifting, and arm switching (Berezikov, 2011). Therefore, variation on miRNA-binding sites indeed can lead to phenotypic innovation, as exemplified by the lineage diversification of cichlid fishes (Loh et al., 2011; Franchini et al., 2016). On the other hand, as target mRNAs can be regulated subtly by several miRNAs, detecting phenotypical effects by population variation seems to be harder than for genetic variation in regulatory or coding regions. Indeed, most of single nucleotide polymorphisms (SNPs) involved in the creation of novel miRNA target sites does not correlate with phenotypic differences among humans (Saunders et al., 2007).

Nonetheless, it is possible that genomic comparisons of different individuals can give insights on the origination process of miRNAs, as previously done for other genetic elements (Hatcher, 2000; Schlötterer, 2015). For instance, the basis of retrogene origination in metazoans has been recently deciphered through *Drosophila* population data. Flanking regions signatures of polymorphic retrocopies revealed that long terminal repeat (LTR) retrotransposons have mediate their formation (Tan et al., 2016). miRNAs are mostly originated *de novo* or by duplication (Meunier et al., 2013), but mechanistic details on how those processes occur are still unknown. Population genomics might help uncover those components through the identification of mutational signatures attached to polymorphic miRNAs that are usually erased by time and throughout their fixation.

In addition, comparing fixed patterns present in different species to polymorphic states observed in a group of individuals are useful tools for contrasting genomic features driven by natural selection to patterns produced by mutation bias (Long et al., 2013). Notable, this type of comparison helped to support the hypothesis in which natural selection drives retrogene duplication from the X chromosome to the autosomes in *Drosophila* and humans (Schridder et al., 2011, 2013; Navarro and Galante, 2015). Therefore, the analyses of different human populations can give further support to the adapted

pattern of miRNAs organized inside old protein coding host genes.

Furthermore, as miRNAs expression and targeting has been shown to be implicated in a wide of human diseases (Mendell and Olson, 2012), seed, and mature region variants found among ethnic populations become clinically important (Rawlings-Goss et al., 2014). More specifically, there are distinct miRNA profiles in diseases between African and European descendants (e.g., Huang et al., 2011; Heegaard et al., 2012) which could be responsible for differences among those populations in susceptibility to diseases, drug sensitiveness, and biomarker diagnostics (Rawlings-Goss et al., 2014). Therefore, should worth investigating if ethnic group variation on miRNA genomic context have also significant role in human health.

From the discussion above, it turns out that the genomic context, as an outcome of natural selection, imposes evolutionary constraints to maintain the structural and functional integrity of its genetic elements. Moreover, it can also propel the evolutionary fate of new elements that arise in a suitable environment, eventually accelerating the process of functionalization. Therefore, evolutionary models tackling the 3D chromatin organization will be of extreme value to pursue the general principles that afford those processes take place throughout genomes.

AUTHOR CONTRIBUTIONS

GSF and MDV conceived the study. GSF, LCH, PAFG, and MDV wrote the manuscript. All authors revised and approved it for publication.

ACKNOWLEDGMENTS

This work was supported by CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) (381128/2016-4, to G.S.F) and FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) (15/20844-4, to M.D.V).

REFERENCES

- Ason, B., Darnell, D. K., Wittbrodt, B., Berezikov, E., Kloosterman, W. P., Wittbrodt, J., et al. (2006). Differences in vertebrate microRNA expression. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14385–14389. doi: 10.1073/pnas.0603529103
- Barbash, S., Shifman, S., and Soreq, H. (2014). Global coevolution of human micromRNAs and their target genes. *Mol. Biol. Evol.* 31, 1237–1247. doi: 10.1093/molbev/msu090
- Barrón, M. G., Fiston-Lavier, A.-S., Petrov, D. A., and González, J. (2014). Population genomics of transposable elements in *Drosophila*. *Annu. Rev. Genet.* 48, 561–581. doi: 10.1146/annurev-genet-120213-092359
- Baskerville, S., and Bartel, D. P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* 11, 241–247. doi: 10.1261/rna.7240905
- Berezikov, E. (2011). Evolution of microRNA diversity and regulation in animals. *Nat. Rev. Genet.* 12, 846–860. doi: 10.1038/nrg3079
- Berezikov, E., Thuemmler, F., van Laake, L. W., Kondova, I., Bontrop, R., Cuppen, E., et al. (2006). Diversity of microRNAs in human and chimpanzee brain. *Nat. Genet.* 38, 1375–1377. doi: 10.1038/ng1914
- Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csárdi, G., Harrigan, P., et al. (2011). The evolution of gene expression levels in mammalian organs. *Nature* 478, 343–348. doi: 10.1038/nature10532
- Brodde, O.-E., and Michel, M. C. (1999). Adrenergic and muscarinic receptors in the human heart. *Pharmacol. Rev.* 51, 651–690.
- Campo-Paysaa, F., Sémon, M., Cameron, R. A., Peterson, K. J., and Schubert, M. (2011). microRNA complements in deuterostomes: origin and evolution of microRNAs. *Evol. Dev.* 13, 15–27. doi: 10.1111/j.1525-142X.2010.00452.x
- Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., et al. (2005). The transcriptional landscape of the mammalian genome. *Science* 309, 1559–1563. doi: 10.1126/science.1112014
- Caron, H., van Schaik, B., van der Mee, M., Baas, F., Riggins, G., van Sluis, P., et al. (2001). The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science* 291, 1289–1292. doi: 10.1126/science.1056794
- Chen, K., and Rajewsky, N. (2007). The evolution of gene regulation by transcription factors and microRNAs. *Nat. Rev. Genet.* 8, 93–103. doi: 10.1038/nrg1990

- Chiang, H. R., Schoenfeld, L. W., Ruby, J. G., Auyeng, V. C., Spies, N., Baek, D., et al. (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev.* 24, 992–1009. doi: 10.1101/gad.1884710
- Cooley, N., Cowley, M. J., Lin, R. C. Y., Marasco, S., Wong, C., Kaye, D. M., et al. (2012). Influence of atrial fibrillation on microRNA expression profiles in left and right atria from patients with valvular heart disease. *Physiol. Genomics* 44, 211–219. doi: 10.1152/physiolgenomics.00111.2011
- Costache, M., Apoil, P. A., Cailleau, A., Elmgren, A., Larson, G., Henry, S., et al. (1997). Evolution of fucosyltransferase genes in vertebrates. *J. Biol. Chem.* 272, 29721–29728. doi: 10.1074/jbc.272.47.29721
- França, G. S., Vibranovski, M. D., and Galante, P. A. F. (2016). Host gene constraints and genomic context impact the expression and evolution of human microRNAs. *Nat. Commun.* 7:11438. doi: 10.1038/ncomms11438
- Franchini, P., Xiong, P., Fruciano, C., and Meyer, A. (2016). The role of microRNAs in the repeated parallel diversification of lineages of midas cichlid fish from nicaragua. *Genome Biol. Evol.* 8, 1543–1555. doi: 10.1093/gbe/evw097
- Fromm, B., Billipp, T., Peck, L. E., Johansen, M., Tarver, J. E., King, B. L., et al. (2015). A Uniform system for the annotation of vertebrate microRNA genes and the evolution of the human microRNAome. *Annu. Rev. Genet.* 49, 213–242. doi: 10.1146/annurev-genet-120213-092023
- Galante, P. A. F., Vidal, D. O., de Souza, J. E., Camargo, A. A., and de Souza, S. J. (2007). Sense-antisense pairs in mammals: functional and evolutionary considerations. *Genome Biol.* 8:R40. doi: 10.1186/gb-2007-8-3-r40
- Gallego, A., Melé, M., Balcells, L., García-Ramallo, E., Torruella-Loran, I., Fernández-Bellón, H., et al. (2016). Functional implications of human-specific changes in great ape microRNAs. *PLoS ONE* 11:e0154194. doi: 10.1371/journal.pone.0154194
- Ghanbarian, A. T., and Hurst, L. D. (2015). Neighboring genes show correlated evolution in gene expression. *Mol. Biol. Evol.* 32, 1748–1766. doi: 10.1093/molbev/msv053
- Graur, D., Zheng, Y., and Azevedo, R. B. (2015). An evolutionary classification of genomic function. *Genome Biol. Evol.* 7, 642–645. doi: 10.1093/gbe/evv021
- Hatcher, M. J. (2000). Persistence of selfish genetic elements: population structure and conflict. *Trends Ecol. Evol.* 15, 271–277. doi: 10.1016/S0169-5347(00)01875-9
- Hausser, J., and Zavolan, M. (2014). Identification and consequences of miRNA-target interactions—beyond repression of gene expression. *Nat. Rev. Genet.* 15, 599–612. doi: 10.1038/nrg3765
- Heegaard, N. H., Schetter, A. J., Welsh, J. A., Yoneda, M., Bowman, E. D., and Harris, C. C. (2012). Circulating micro-RNA expression profiles in early stage nonsmall cell lung cancer. *Int. J. Cancer* 130, 1378–1386. doi: 10.1002/ijc.26153
- Hinske, L. C., França, G. S., Torres, H. A. M., Ohara, D. T., Lopes-Ramos, C. M., Heyn, J., et al. (2014). miRIAD-integrating microRNA inter- and intragenic data. *Database (Oxford)*. 2014, 1–9. doi: 10.1093/database/bau099
- Hinske, L. C. G., Galante, P. A. F., Kuo, W. P., and Ohno-Machado, L. (2010). A potential role for intragenic miRNAs on their hosts' interactome. *BMC Genomics* 11:533. doi: 10.1186/1471-2164-11-533
- Huan, T., Rong, J., Liu, C., Zhang, X., Tanriverdi, K., Joehanes, R., et al. (2015). Genome-wide identification of microRNA expression quantitative trait loci. *Nat. Commun.* 6:6601. doi: 10.1038/ncomms7601
- Huang, R. S., Gamazon, E. R., Wen, Y., Im, H. K., Zhang, W., Wing, C., et al. (2011). Population differences in microRNA expression and biological implications. *RNA Biol.* 8, 692–701. doi: 10.4161/rna.8.4.16029
- Iwama, H., Kato, K., Imachi, H., Murao, K., and Masaki, T. (2013). Human microRNAs originated from two periods at accelerated rates in mammalian evolution. *Mol. Biol. Evol.* 30, 613–626. doi: 10.1093/molbev/mss262
- Juliant, S., Harduin-Lepers, A., Monjaret, F., Catieau, B., Violet, M. L., Cérutti, P., et al. (2014). The $\alpha 1$, 6-fucosyltransferase gene (FUT8) from the Sf9 lepidopteran insect cell line: insights into FUT8 evolution. *PLoS ONE* 9:e110422. doi: 10.1371/journal.pone.0110422
- Kapranov, P., Willingham, A. T., and Gingeras, T. R. (2007). Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* 8, 413–423. doi: 10.1038/nrg2083
- Liu, Y.-C., Yen, H.-Y., Chen, C.-Y., Chen, C.-H., Cheng, P.-F., Juan, Y.-H., et al. (2011). Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11332–11337. doi: 10.1073/pnas.1107385108
- Loh, Y. E., Yi, S. V., and Strelman, J. T. (2011). Evolution of microRNAs and the diversification of species. *Genome Biol. Evol.* 3, 55–65. doi: 10.1093/gbe/evq085
- Long, M., VanKuren, N. W., Chen, S., and Vibranovski, M. D. (2013). New gene evolution: little did we know. *Annu. Rev. Genet.* 47, 307–333. doi: 10.1146/annurev-genet-111212-133301
- Lyu, Y., Shen, Y., Li, H., Chen, Y., Guo, L., Zhao, Y., et al. (2014). New microRNAs in *Drosophila*—birth, death and cycles of adaptive evolution. *PLoS Genet.* 10:e1004096. doi: 10.1371/journal.pgen.1004096
- Marsico, A., Huska, M. R., Lasserre, J., Hu, H., Vucicevic, D., Musahl, A., et al. (2013). PROMiRNA: a new miRNA promoter recognition method uncovers the complex regulation of intronic miRNAs. *Genome Biol.* 14:R84. doi: 10.1186/gb-2013-14-8-r84
- McLoughlin, H. S., Wan, J., Spengler, R. M., Xing, Y., and Davidson, B. L. (2014). Human-specific microRNA regulation of FOXO1: implications for microRNA recognition element evolution. *Hum. Mol. Genet.* 23, 2593–2603. doi: 10.1093/hmg/ddt655
- Mele, M., Ferreira, P. G., Reverter, F., DeLuca, D. S., Monlong, J., Sammeth, M., et al. (2015). The human transcriptome across tissues and individuals. *Science* 348, 660–665. doi: 10.1126/science.aaa0355
- Mellor, J., Woloszczuk, R., and Howe, F. S. (2016). The interleaved genome. *Trends Genet.* 32, 57–71. doi: 10.1016/j.tig.2015.10.006
- Mendell, J. T., and Olson, E. N. (2012). MicroRNAs in stress signaling and human disease. *Cell* 148, 1172–1187. doi: 10.1016/j.cell.2012.02.005
- Mercer, T. R., and Mattick, J. S. (2013). Understanding the regulatory and transcriptional complexity of the genome through structure. *Genome Res.* 23, 1081–1088. doi: 10.1101/gr.156612.113
- Meunier, J., Lemoine, F., Soumillon, M., Liechti, A., Weier, M., Guschanski, K., et al. (2013). Birth and expression evolution of mammalian microRNA genes. *Genome Res.* 23, 34–45. doi: 10.1101/gr.140269.112
- Michalak, P. (2008). Coexpression, coregulation, and cofunctionality of neighboring genes in eukaryotic genomes. *Genomics* 91, 243–248. doi: 10.1016/j.ygeno.2007.11.002
- Mlynárová, L., Loonen, A., Mietkiewska, E., Jansen, R. C., and Nap, J. P. (2002). Assembly of two transgenes in an artificial chromatin domain gives highly coordinated expression in tobacco. *Genetics* 160, 727–740.
- Navarro, F. C. P., and Galante, P. A. F. (2015). A genome-wide landscape of retrocopies in primate genomes. *Genome Biol. Evol.* 7, 2265–2275. doi: 10.1093/gbe/evv142
- Ninova, M., Ronshaugen, M., and Griffiths-Jones, S. (2014). Fast-evolving microRNAs are highly expressed in the early embryo of *Drosophila virilis*. *RNA* 20, 360–372. doi: 10.1261/rna.041657.113
- Nozawa, M., Fujimi, M., Iwamoto, C., Onizuka, K., Fukuda, N., Ikeo, K., et al. (2016). Evolutionary transitions of MicroRNA-target pairs. *Genome Biol. Evol.* 8, 1621–1633. doi: 10.1093/gbe/evw092
- Ozsolak, F., Poling, L. L., Wang, Z., Liu, H., Liu, X. S., Roeder, R. G., et al. (2008). Chromatin structure analyses identify miRNA promoters Chromatin structure analyses identify miRNA promoters. *Genes Dev.* 22, 3172–3183. doi: 10.1101/gad.1706508
- Park, S. G., and Choi, S. S. (2010). Expression breadth and expression abundance behave differently in correlations with evolutionary rates. *BMC Evol. Biol.* 10:241. doi: 10.1186/1471-2148-10-241
- Pelechano, V., and Steinmetz, L. M. (2013). Gene regulation by antisense transcription. *Nat. Rev. Genet.* 14, 880–893. doi: 10.1038/nrg3594
- Purmann, A., Toedling, J., Schueler, M., Carninci, P., Lehrach, H., Hayashizaki, Y., et al. (2007). Genomic organization of transcriptomes in mammals: coregulation and cofunctionality. *Genomics* 89, 580–587. doi: 10.1016/j.ygeno.2007.01.010
- Rasmussen, M. H., Lyskjær, I., Jersie-Christensen, R. R., Tarpgaard, L. S., Primdal-Bengtson, B., Nielsen, M. M., et al. (2016). miR-625-3p regulates oxaliplatin resistance by targeting MAP2K6-p38 signalling in human colorectal adenocarcinoma cells. *Nat. Commun.* 7:12436. doi: 10.1038/ncomms12436
- Rawlings-Goss, R. A., Campbell, M. C., and Tishkoff, S. A. (2014). Global population-specific variation in miRNA associated with cancer risk and clinical biomarkers. *BMC Med. Genomics* 7:53. doi: 10.1186/1755-8794-7-53
- Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. *Genome Res.* 14, 1902–1910. doi: 10.1101/gr.2722704

- Saunders, M. A., Liang, H., and Li, W. H. (2007). Human polymorphism at microRNAs and microRNA target sites. *Proc. Natl. Acad. Sci. U.S.A.* 104, 3300–3305. doi: 10.1073/pnas.0611347104
- Schlötterer, C. (2015). Genes from scratch—the evolutionary fate of de novo genes. *Trends Genet.* 31, 215–219. doi: 10.1016/j.tig.2015.02.007
- Schrider, D. R., Navarro, F. C. P., Galante, P. A. F., Parmigiani, R. B., Camargo, A. A., Hahn, M. W., et al. (2013). Gene copy-number polymorphism caused by retrotransposition in humans. *PLoS Genet.* 9:e1003242. doi: 10.1371/journal.pgen.1003242
- Schrider, D. R., Stevens, K., Cardeño, C. M., Langley, C. H., and Hahn, M. W. (2011). Genome-wide analysis of retrogene polymorphisms in *Drosophila melanogaster*. *Genome Res.* 21, 2087–2095. doi: 10.1101/gr.116434.110
- Shen, J., Xiao, Z., Wu, W. K. K., Wang, M. H., To, K. F., Chen, Y., et al. (2015). Epigenetic silencing of miR-490-3p reactivates the chromatin remodeler SMARCD1 to promote *Helicobacter pylori*-induced gastric carcinogenesis. *Cancer Res.* 75, 754–765. doi: 10.1158/0008-5472.CAN-14-1301
- Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., et al. (2015). An integrated map of structural variation in 2,504 human genomes. *Nature* 526, 75–81. doi: 10.1038/nature15394
- Sun, Y., Chen, D., Cao, L., Zhang, R., Zhou, J., Chen, H., et al. (2013). MiR-490-3p modulates the proliferation of vascular smooth muscle cells induced by ox-LDL through targeting PAPP-A. *Cardiovasc. Res.* 100, 272–279. doi: 10.1093/cvr/cvt172
- Tan, S., Cardoso-Moreira, M., Shi, W., Zhang, D., Huang, J., Mao, Y., et al. (2016). LTR-mediated retroposition as a mechanism of RNA-based duplication in metazoans. *Genome Res.* 26, 1–13. doi: 10.1101/gr.204925.116
- Taylor, R. S., Tarver, J. E., Hiscock, S. J., and Donoghue, P. C. J. (2014). Evolutionary history of plant microRNAs. *Trends Plant Sci.* 19, 175–182. doi: 10.1016/j.tplants.2013.11.008
- Verma, A. M., Patel, M., Aslam, M. I., Jameson, J., Pringle, J. H., Wurm, P., et al. (2015). Circulating plasma microRNAs as a screening method for detection of colorectal adenomas. *Lancet* 385, S100. doi: 10.1016/S0140-6736(15)60415-9
- Wang, X., Inoue, S., Gu, J., Miyoshi, E., Noda, K., Li, W., et al. (2005). Dysregulation of TGF- β 1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15791–15796. doi: 10.1073/pnas.0507375102
- Wang, Y., Luo, J., Zhang, H., and Lu, J. (2016). microRNAs in the same clusters evolve to coordinately regulate functionally related genes. *Mol. Biol. Evol.* 33, 2232–2247. doi: 10.1093/molbev/msw089
- Wheeler, B. M., Heimberg, A. M., Moy, V. N., Sperling, E. A., Holstein, T. W., Heber, S., et al. (2009). The deep evolution of metazoan microRNAs. *Evol. Dev.* 11, 50–68. doi: 10.1111/j.1525-142X.2008.00302.x
- Wolf, Y. I., Novichkov, P. S., Karev, G. P., Koonin, E. V., and Lipman, D. J. (2009). The universal distribution of evolutionary rates of genes and distinct characteristics of eukaryotic genes of different apparent ages. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7273–7280. doi: 10.1073/pnas.0901808106
- Xu, J., Zhang, R., Shen, Y., Liu, G., Lu, X., and Wu, C. -I. (2013). The evolution of evolvability in microRNA target sites in vertebrates. *Genome Res.* 23, 1810–1816. doi: 10.1101/gr.148916.112
- Zhang, L., Hu, A., Yuan, H., Cui, L., Miao, G., Yang, X., et al. (2008). A missense mutation in the CHRM2 gene is associated with familial dilated cardiomyopathy. *Circ. Res.* 102, 1426–1432. doi: 10.1161/CIRCRESAHA.107.167783
- Zheng, H., Ma, R., Wang, Q., Zhang, P., Li, D., Wang, Q., et al. (2015). MiR-625-3p promotes cell migration and invasion via inhibition of SCAI in colorectal carcinoma cells. *Oncotarget* 6, 27805–27815. doi: 10.18632/oncotarget.4738
- Zhou, X., Zhang, C. Z., Lu, S. -X., Chen, G. G., Li, L. -Z., Liu, L. -L., et al. (2014). miR-625 suppresses tumour migration and invasion by targeting IGF2BP1 in hepatocellular carcinoma. *Oncogene* 34, 965–977. doi: 10.1038/onc.2014.35

Conflict of Interest Statement: 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 © 2017 França, Hinske, Galante and Vibrantovski. 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) or licensor 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.