HHS Public Access

Author manuscript

J Mol Evol. Author manuscript; available in PMC 2016 April 01.

Published in final edited form as:

J Mol Evol. 2015 April ; 80(0): 186–188. doi:10.1007/s00239-015-9668-x.

The “life histories” of genes

Esther Betrán*

Department of Biology, University of Texas at Arlington

Genes do not only originate. Genes have their own life histories. Life histories of genes might differ in the way the genes originate, but also in how fast they become established, and how long they persist in the genome. To understand the diversity of life histories of genes it is necessary to consider genes of various ages as each age group can illustrate a different kind of life history. We can find examples of the “live fast, die young” existence only among young gene cohorts and of the “live slow, die old” lifestyle within old genes. We gain this information using the comparative genomics approach that is becoming increasingly productive thanks to the abundance of whole-genome sequence data generated within the last 20 years. Often, this genomic information is accompanied by age-, tissue- and sex-specific gene expression data as well as functional studies that can give us a sense of the gene’s function/sand complement the evolutionary analyses.

There are multiple mechanisms for the origin of genes including gene duplication, horizontal gene transfer, domestication of transposable elements or viruses, and *de novo* formation from non-coding sequences. These mechanisms generate new protein-coding genes as well as new non-coding RNA genes (e.g., microRNAs) and might produce duplications at different rates in different regions of the genome or in different lineages. Most of these new genes are lost without ever reaching fixation in a population. Many of those that do fix are pseudogenized in the absence of selection for their retention. But a small fraction of new genes is functional and is integrated into previously established gene networks that participate in various biological processes. The differences in the life histories of these new genes include the origination process, how fast they become established in the genome, the strength and nature of selection they experience, and their life span. I will highlight some contrasting life histories here, although there are many more possibilities that fall in between these. I will argue that thinking about the genes from the perspective of their life histories helps us recognize gene turnover patterns (i.e., patterns of recurrent gene gain and gene loss) and, consequently, should help us understand the selective pressures experienced by diverse tissues and pathways in different lineages.

Many genes are born through gene/genome duplication or gene recombination (i.e., from preexisting genes). These are the new genes we currently know the most about and include the best instances of the differing life histories. Some developmental regulatory genes constitute examples of long-lived, conserved gene duplications. For example, there have been expansions of *Hox* genes through tandem duplication that account for changes in body

*Corresponding address: Esther Betrán, Biology Department, Box 19498, University of Texas at Arlington, Arlington, TX 76019, USA, Phone (817) 272 1446; Fax (817) 272 2855; betran@uta.edu.

plans by providing diverged transcription factors that specify the identity of different segments. Despite some turnover or additional duplications in some lineages, some *Hox* gene homeodomains have remained conserved and their biochemical functions have remained the same despite regulating different sets of genes (e.g., overexpression of *Hoxb1*, the vertebrate ortholog of the fly *labial* gene, or of *labial* itself in flies show similar phenotypes). Such conservation implies strong purifying selection which in turn suggests that the process of development is not easily modified. *Hox* genes are transcription factors that regulate many downstream genes and changes in their DNA-binding domain are likely to have disastrous consequences for development.

For a younger gene cohort, it has been observed that some relatively recently duplicated genes have become essential. In particular, Chen et al. (2010) targeted young fly genes (most of them gene duplicates) using RNAi and found that ~30 % have become essential for viability. Many of these genes are expressed in late larva and their knockdown leads to pupae arrest. These developmental stages are known to transcribe genes of intermediate age (i.e., these tissues experience an intermediate level of turnover; Domazet-Loso et al. 2010). While it is still unknown how many of these genes are essential because they partition the function of an essential gene, there are examples showing how such genes can become essential by acquiring new functions. In one instance, the gene *Umbrea* acquired essential centromeric function owing to a loss of a heterochromatin-binding domain and several changes in its amino acid sequence while the loss of function of the parental gene encoding a heterochromatin protein does not compromise viability (Ross et al. 2013). This neofunctionalization to an essential function must have contributed to increased life span of *Umbrea* in *Drosophila melanogaster* as this gene is lost in some lineages where such changes did not occur. Ross and colleagues propose that genetic conflict involving centromeric function might require the recurrent recruitment of new proteins. This example reveals an essential pathway that actually experiences some gene turnover.

Among even younger genes, we now have ample evidence that some duplicated genes are under strong selection for high turnover and for frequent changes in the protein sequence. Such genes are often involved in interactions with the environment or participate in arms races, including male-male competition, male-female antagonism, and host defense against infections or selfish genetic elements. The life histories of these genes can be illustrated by some testis-specific genes in Drosophila. In some instances selection for male-specific functions has been proposed to act even before the gene is duplicated, generating balanced polymorphism/allelic divergence at the parental gene (Connallon and Clark 2011; Gallach and Betran 2011). Subsequent gene duplication may involve gene relocation that facilitates the acquisition of a male germline specific expression pattern and prevents gene conversion (Gallach and Betran 2011; Sorourian et al. 2014) enabling the evolution of a new function. This explains why under such circumstances it is often the new copy that acquires a new function. Many such genes do not remain in the genome for long, but are quickly lost, presumably because their function becomes obsolete in this fast evolving tissue. This might explain why many young genes are expressed in testis but not as many older genes are transcribed in this tissue providing an additional explanation to previous observations by Vinckenbosch et al. (2006). Drosophila testis is thus a dynamic tissue where proteins evolve

fast, where recruitment of new genes occurs frequently and where gene loss is also common. Consistent with this, the adult males have been shown to express younger genes than adult females or other life cycle stages (Domazet-Loso et al. 2010).

The high turnover tissue or stage does not always need to be an adult male tissue. In zebrafish, early embryogenesis appears to transcribe a lot of young genes (Domazet-Loso et al. 2010), although adult male tissues still transcribe a lot of young genes in zebrafish. High turnover of some non-coding RNA genes has also been observed. Adult male tissues show expression of these young non-coding genes (Lyu et al. 2014) in *D. melanogaster*, but turnover is observed during early embryogenesis for microRNAs in *D. virilis* (Ninova et al. 2014). Thus, the prevalence of particular gene life histories differs between tissues, protein-coding and non-coding genes and between lineages as well.

There are many other variations in such “live fast, die young” gene life history. For example, in the absence of relocation of the duplicate, gene conversion can speed up adaptation for the whole gene family by facilitating the exploration of adaptive combinations through ectopic recombination. That is, some fast evolving/fast turnover genes are organized in tandemly duplicated gene families (Kelleher and Markow 2009). It has also been proposed that in some instances the fast turnover might occur even if a beneficial dose effect is the important attribute of the new gene, e.g., the duplication of some genes believed to be suppressors of selfish male meiotic drive systems (Phadnis et al. 2012).

Why do these differences in life histories exist? Our current knowledge of gene functions, biological processes they participate in, and their patterns of molecular evolution builds a multi-layered, telescoping image of an organism. There are processes that form the basic cellular and developmental pathways of a functioning organism, maintain homeostasis and create a stable environment that allows for other, more dynamic processes. Selection pressures dictate how dynamic the processes are, and the life histories of genes reflect the process they are recruited into. Conserved, long-lived genes are frequently involved in fundamental biological processes that are rarely disturbed and that recruit new genes less often, while fast-evolving genes that are prone to loss and recurrent duplication participate in the processes that are under strong selection for change.

In addition to gene birth through duplication, other processes like horizontal gene transfer and transposable element or virus domestication are known to give rise to new genes. Somewhat unexpectedly, data are accumulating to support *de novo* gene birth as well. New genes that arise through these mechanisms differ from genes that are born from pre-existing genes in one important aspect: they are foreign to the host genome. What life histories will they have? What tissues are going to recruit these foreign new genes? One clear prediction is that tissues that are enriched for dynamic and fast-evolving processes would be more likely to incorporate these foreign genes. In contrast, limited recruitment of these genes is expected in tissues with highly constrained pathways. Is this what has been observed? Some data has accumulated recently for *de novo* genes and it seems that, yes, again *de novo* genes are often transcribed and acquire new functions in Drosophila and human testis (Xie et al. 2012; Zhao et al. 2014; Palmieri et al. 2014) and in human brain (another tissue that shows quite a bit of turnover in humans; Xie et al. 2012). Tissues under strong pressures to change will

frequently incorporate foreign genes but, in these tissues, new genes should also suffer frequent losses. At the same time, some genes might evolve interactions that could make them survive longer and eventually even express in other tissues. Among genes domesticated from transposable elements or viruses, some are recruited into tissues with conflicts that can lead to high gene turnover. These include placenta where there is a mother-offspring conflict (Malik 2012), or immune cells (Malik and Henikoff 2005). On the other hand, many domesticated genes from transposable elements are old transposases involved in chromatin remodeling or act as transcription factors (Feschotte and Pritham 2007) indicating that they are currently recruited into stable gene networks. As incorporation of a new foreign gene directly into well-functioning pathway is unlikely, these transposases might illustrate an alternative way of establishing stable interactions. Initially, these proteins could have originated as chromatin remodeling proteins within a fast-evolving network (Levine and Malik 2013), but later evolved into transcription factors participating in a stable and long-lasting biological process. This is likely a general route on the way to establishing a permanent residence in the genome. New genes are most often recruited to carry out short-lived functions in highly dynamic tissues, but occasionally secure additional, more central roles.

The study of genes in the context of their life histories reveals the extent of gene turnover making it necessary to consider the biological processes the genes are involved in and forcing us to examine the precise selective pressures that govern the evolution of these processes.

Acknowledgments

I thank Anna Williford for numerous discussions and Anna Williford and Jeff Demuth for reading and providing comments to the text. The Betrán laboratory is supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R01GM071813 to E.B. The content of this publication is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health.

References

Chen S, Zhang YE, Long M. New genes in Drosophila quickly become essential. Science. 2010; 330(6011):1682–1685. [PubMed: 21164016]

Connallon T, Clark AG. The Resolution of Sexual Antagonism by Gene Duplication. Genetics. 2011; 187(3):919–937. [PubMed: 21220356]

Domazet-Loso T, Tautz D. A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. Nature. 2010; 468(7325):815–818. [PubMed: 21150997]

Feschotte C, Pritham EJ. DNA transposons and the evolution of eukaryotic genomes. Annu Rev Genet. 2007; 41:331–368. [PubMed: 18076328]

Gallach M, Betran E. Intralocus sexual conflict resolved through gene duplication. Trends Ecol Evol. 2011; 26(5):222–228. [PubMed: 21397976]

Kelleher ES, Markow TA. Duplication, selection and gene conversion in a Drosophila mojavensis female reproductive protein family. Genetics. 2009; 181(4):1451–1465. [PubMed: 19204376]

Levine MT, Malik HS. A rapidly evolving genomic toolkit for Drosophila heterochromatin. Fly (Austin). 2013; 7(3):137–141. [PubMed: 23519206]

Lyu Y, Shen Y, Li H, Chen Y, Guo L, Zhao Y, Hungate E, Shi S, Wu CI, Tang T. New microRNAs in Drosophila--birth, death and cycles of adaptive evolution. PLoS Genet. 2014; 10(1):e1004096. [PubMed: 24465220]

Malik HS. Retroviruses push the envelope for mammalian placentation. Proc Natl Acad Sci U S A. 2012; 109(7):2184–2185. [PubMed: 22308481]

Malik HS, Henikoff S. Positive selection of Iris, a retroviral envelope-derived host gene in Drosophila melanogaster. PLoS Genet. 2005; 1(4):e44. [PubMed: 16244705]

Ninova M, Ronshaugen M, Griffiths-Jones S. Fast-evolving microRNAs are highly expressed in the early embryo of Drosophila virilis. Rna. 2014; 20(3):360–372. [PubMed: 24448446]

Palmieri N, Kosiol C, Schlotterer C. The life cycle of Drosophila orphan genes. eLife. 2014; 3:e01311. [PubMed: 24554240]

Phadnis N, Hsieh E, Malik HS. Birth, death, and replacement of karyopherins in Drosophila. Mol Biol Evol. 2012; 29(5):1429–1440. [PubMed: 22160828]

Ross BD, Rosin L, Thomae AW, Hiatt MA, Vermaak D, de la Cruz AF, Imhof A, Mellone BG, Malik HS. Stepwise evolution of essential centromere function in a Drosophila neogene. Science. 2013; 340(6137):1211–1214. [PubMed: 23744945]

Sorourian M, Kunte MM, Domingues S, Gallach M, Özdil F, Río J, EB. Relocation facilitates the acquisition of short cis-regulatory regions that drive the expression of retrogenes during spermatogenesis in Drosophila. Mol Biol Evol. 2014; 31(8):2170–2180. [PubMed: 24855141]

Vinckenbosch N, Dupanloup I, Kaessmann H. Evolutionary fate of retroposed gene copies in the human genome. Proc Natl Acad Sci U S A. 2006; 103(9):3220–3225. [PubMed: 16492757]

Xie C, Zhang YE, Chen JY, Liu CJ, Zhou WZ, Li Y, Zhang M, Zhang R, Wei L, Li CY. Hominoid-specific de novo protein-coding genes originating from long non-coding RNAs. PLoS Genet. 2012; 8(9):e1002942. [PubMed: 23028352]

Zhao L, Saelao P, Jones CD, Begun DJ. Origin and spread of de novo genes in Drosophila melanogaster populations. Science. 2014; 343(6172):769–772. [PubMed: 24457212]