

## NEWS AND COMMENTARY

### Genome evolution

# Gene duplication and the resolution of adaptive conflict

JF Storz

*Heredity* (2009) 102, 99–100; doi:10.1038/hdy.2008.114; published online 29 October 2008

Gene duplication is known to play an important role in the evolution of novel protein functions. However, there is still much debate about the evolutionary mechanisms that are responsible for the initial retention and subsequent divergence of newly created gene duplicates (Lynch *et al.*, 2001; Zhang, 2003; Lynch and Katju, 2004; Taylor and Raes, 2004). According to the neofunctionalization (NEO-F) model of Ohno (1970), the functional redundancy of duplicated genes entails a relaxation of purifying selection that results in the accumulation of degenerative mutations in one gene copy, while the other duplicate copy continues to perform the essential tasks of the ancestral, single-copy gene. In the vast majority of cases, the redundant gene duplicate will eventually be rendered functionless by inactivating mutations. In a very small minority of cases, the redundant gene may escape this fate by fixing one or more mutations that fortuitously adapt the encoded protein to a new or modified function.

An alternative model for the evolution of novel protein functions, originally suggested by Piatigorsky and Wistow (1991) and Hughes (1994), invokes the presence of functional divergence before gene duplication and does not require a postduplication phase of relaxed functional constraint. This model envisions an ancestral, single-copy gene that encodes a generalist protein that is capable of performing two or more distinct subfunctions. The product of this gene experiences 'adaptive conflict', as joint optimization of the protein's multiple subfunctions is constrained by antagonistic pleiotropy. For example, imagine a 'promiscuous' enzyme that acts on a broad range of different substrates. Newly arisen mutations that increase enzyme activity on substrate A may be prevented from going to fixation because they compromise the enzyme's activity on substrate B. Likewise, mutations that increase enzyme activity on substrate B may compromise the activity on substrate

A. If the underlying gene is duplicated, then each of the two nascent paralogs can break free of these pleiotropic constraints and specialize on a more narrow range of substrates. In this way, gene duplication can resolve adaptive conflicts between competing subfunctions of a pleiotropically constrained single-copy gene. The NEO-F model of Ohno (1970) and the escape from adaptive conflict (EAC) model both invoke the accumulation of mutations that would have been off-limits before duplication. One key distinction is that in the EAC model, the division of labor between the duplicated genes is brought about by the fixation of advantageous mutations that refine or elaborate ancestral subfunctions of the encoded protein.

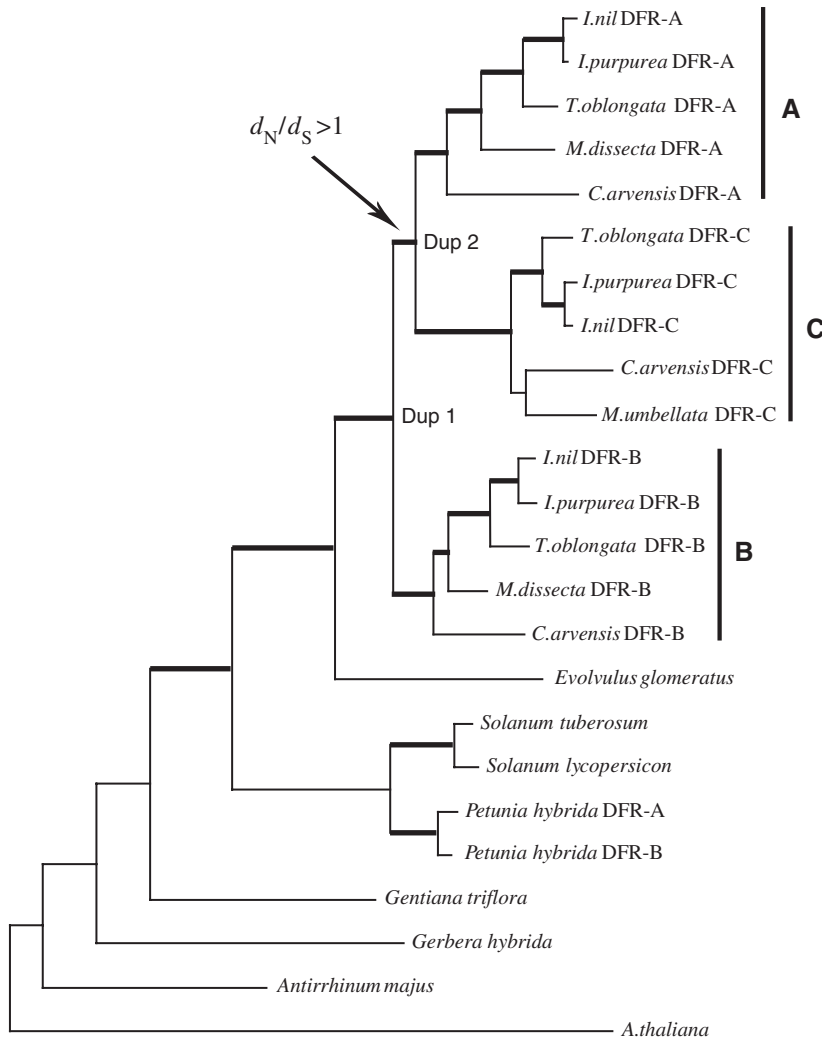
In a recent study of duplicated genes in the anthocyanin biosynthetic pathway of morning glories, Des Marais and Rausher (2008) argue that the evolution of novel enzymatic functions is best explained by the EAC model. The authors reached this conclusion by performing phylogenetically based tests of positive selection in conjunction with experimental assays of enzyme function. The authors used two criteria to distinguish the NEO-F and EAC models. First, they reasoned that EAC entails adaptive change in both duplicate gene copies, whereas NEO-F is expected to involve adaptive changes in only one copy because purifying selection acts to maintain the ancestral function of the other duplicate copy. Second, EAC entails an adaptive improvement of both ancestral and derived protein functions, whereas NEO-F does not necessarily involve any modification of the ancestral function. The authors applied both criteria to distinguish between NEO-F and EAC as explanations for adaptive change in duplicate copies of the anthocyanin biosynthetic pathway gene, dihydroflavonol-4-reductase (DFR).

The DFR enzyme reduces the flavonoid precursors of anthocyanin pigments, and therefore plays a role in producing the red, purple and blue

flower petals of morning glories and other angiosperm plants. A number of morning glory species possess triplicated copies of the DFR gene (DFR-A, DFR-B and -DFRC), and the remainder possess a single copy. Phylogenetic reconstructions of the DFR gene family in morning glory species that possess three DFR paralogs show that an initial duplication event gave rise to DFR-B and the proto-DFR-A/C gene, and a subsequent duplication event gave rise to the separate A and C genes (Figure 1).

Subsequent to the first duplication event that gave rise to the DFR-B and proto-DFR-A/C genes, both of the newly created paralogs appear to have evolved a specialization of function. Adaptive change in the DFR-A/C gene is indicated by a highly accelerated rate of amino-acid substitution in the postduplication branch leading to the DFR-A/C clade (Figure 1). Although a similar excess of amino-acid substitution was not observed in the postduplication branch leading to the DFR-B gene, functional assays revealed that the product of this duplicated gene evolved enhanced activity on ancestral flavonoid substrates. These results satisfy the first criterion for the operation of EAC, as adaptive modifications of enzyme function appear to have occurred in both daughter copies of the first duplication (the DFR-B and DFR-A/C genes). As the postduplication functional changes in DFR-B enzyme activity were not accompanied by any statistically significant excess of amino-acid substitutions, it may be that the causative changes in enzyme function were attributable to a small number of amino-acid substitutions with large effects. The postduplication improvement of ancestral protein function in the DFR-B gene also satisfied the second criterion for the operation of EAC. The inference is that the ability of DFR to reduce dihydroflavonols was pleiotropically constrained in the ancestral, single-copy gene, and this constraint was then alleviated by gene duplication.

Des Marais and Rausher (2008) argue that EAC may be a relatively common explanation for the retention and functional divergence of duplicated genes. Other possible cases that have recently been reported include duplicated globin genes of rodents that direct the synthesis of hemoglobin isoforms with graded oxygen-binding affinities and/or auto-oxidation rates (Storz *et al.*, 2008), and duplicated genes in the galactose -use pathway of yeast that have evolved specialized transcriptional



**Figure 1** Gene tree of dihydroflavonol-4-reductase (DFR) coding sequences for morning glory species that possess three DFR gene copies (A, B and C) and out-group species that possess a single copy (or an independently derived pair of duplicated copies in the case of *Petunia hybrida*). The postduplication branch that leads to the progenitor of the DFR-A and DFR-C genes was characterized by a highly elevated rate of amino-acid substitution, as indicated by a  $d_N/d_S$  ratio  $\gg 1$  (where  $d_N$  is the rate of nonsynonymous substitution per nonsynonymous site and  $d_S$  is the rate of synonymous substitution per synonymous site). This lineage-specific excess of amino-acid substitution suggests that the functional divergence of the DFR-B and DFR-A/C gene copies was driven by positive selection that adapted the proto-DFR-A/C isozyme to a new or modified function.

control of coinducer and galactokinase functions (Hittinger and Carroll, 2007). To assess the generality of the EAC model, more studies will be required to determine whether specialization of

function between duplicated genes typically evolves as a refinement of preexisting subfunctions of a pleiotropically constrained ancestral gene.

One of the especially commendable aspects of the study by Des Marais and Rausher (2008) is that molecular evolution analyses were integrated with experimental studies of enzyme function. This study illustrates the growing appreciation that statistical inferences about positive selection based on comparative sequence analysis need to be buttressed by functional data in order to draw firm conclusions about molecular adaptation.

JF Storz is at the School of Biological Sciences, University of Nebraska, Lincoln, NE, 68588, USA.

e-mail: jstorz2@unl.edu

- Des Marais DL, Rausher MD (2008). Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature* **454**: 762–765.
- Hittinger CT, Carroll SB (2007). Gene duplication and the adaptive evolution of a classic genetic switch. *Nature* **449**: 677–682.
- Hughes AL (1994). The evolution of functionally novel proteins after gene duplication. *Proc Roy Soc Lond B* **256**: 119–124.
- Lynch M, O'Hely M, Walsh B, Force A (2001). The probability of preservation of a newly arisen gene duplicate. *Genetics* **159**: 1789–1804.
- Lynch M, Katju V (2004). The altered evolutionary trajectories of gene duplicates. *Trends Genet* **20**: 544–549.
- Ohno S (1970). *Evolution by Gene Duplication*. Springer-Verlag: New York.
- Piatigorsky J, Wistow G (1991). The recruitment of crystallins: new functions precede gene duplication. *Science* **252**: 1078–1079.
- Storz JF, Hoffmann FG, Opazo JC, Moriyama H (2008). Adaptive functional divergence among triplicated  $\alpha$ -globin genes in rodents. *Genetics* **178**: 1623–1638.
- Taylor JS, Raes J (2004). Duplication and divergence: the evolution of new genes and old ideas. *Annu Rev Genet* **38**: 615–643.
- Zhang J (2003). Evolution by gene duplication. *Trends Ecol Evol* **18**: 292–298.

#### Editor's suggested reading

- Sjödin P, Hedman H, Shavorskaya O, Finet C, Lascoux M, Lagercrantz U (2007). Recent degeneration of an old duplicated flowering time gene in *Brassica nigra*. *Heredity* **98**: 375–384.
- Torgerson DG, Singh RS (2006). Enhanced adaptive evolution of sperm-expressed genes on the mammalian X chromosome. *Heredity* **96**: 39–44.
- Van de Peer Y (2006). Evolutionary genetics: when duplicated genes don't stick to the rules. *Heredity* **96**: 204–205.
- Wittkopp PJ (2006). Evolution of cis-regulatory sequence and function in Diptera. *Heredity* **97**: 139–147.