Heredity (2008) 101, 299–300

www.nature.com/hdy

NEWS AND COMMENTARY

Gene duplication

Red queens, linkage, redundancy and synfunctionalization

DA Liberles

Heredity (2008) **101,** 299–300; doi:10.1038/hdy.2008.81; published online 13 August 2008

The interplay between gene duplication, positive selection and shifting gene functions takes many complex turns. In *PLoS Pathogens*, Sawyer *et al*. (2007) report on the intriguing story of duplication in the TRIM genes. TRIM proteins are multidomain proteins with well-documented antiviral activity (see Nisole *et al*., 2005 for a review). They contain a zinc-finger RING domain involved in protein–protein interaction, one to two B-boxes of unknown function, a coiled-coil domain involved in protein homo-oligomerization and a C-terminal SPRY domain that has been proposed to play a role in either protein–protein or protein–RNA interactions. The precise mechanism of viral inhibition is unclear, although hypotheses include perturbation of viral uncoating, physical prevention of interactions between viral and host proteins, a role in viral sequestration within the cell and the possibility of ubiquitin-conjugating activity (Nisole *et al*., 2005). Because of the antiviral activity and duplication, these proteins have evidently been subjected to an interesting pattern of lineage-specific positive selection.

Evidence has accumulated that genes involved in the immune system have evolved rapidly and under positive selection. These data include one of the first comparisons of the ratio of nonsynonymous to synonymous nucleotide substitution rates (dN/dS) by Hughes and Nei (1988, 1989), which suggested that positive selection had acted on both the MHC class I and II genes. Analysis of sequences from the adaptive evolution database has shown that this evolution was sometimes lineage-specific. A systematic calculation of dN/dS ratios across chordate gene families revealed that immune system genes were among those most commonly found to be under positive selection (Liberles *et al*., 2001). These observations are consistent with the view of pathogen–host coevolution driving an arms race for survival that results in particularly fast evolution and positive selection in response to adaptations in the adversary. As neither gains the upper hand, pathogen-host evolution can be described by the 'Red Queen hypothesis': just like Lewis Carol's heroine, Alice, in *Through the Looking-Glass;* they are running (i.e., evolving) as fast as they can, just to stay in the same place. The TRIM genes show evolutionary patterns that could have resulted from such repeated positive selection, and which are lineage-specific, presumably because they have responded to provide defense against specific evolving viral pathogens.

Comparisons between two genes, *TRIM22* and *TRIM5* are particularly interesting. This pair appears to have been produced by an ancient gene duplication in the common ancestor of eutherian mammals, but show complementary, anticorrelated patterns of positive selection in hominoids and Old World monkeys. This pattern sheds light on explanations for the retention of duplicate genes. While geneticists point out the lack of phenotype for many gene knockouts (Cooke *et al*., 1997) as evidence of the redundancy of duplicate genes, evolutionary theory points in exactly the opposite direction. If the redundancy were being maintained by selection, then population genetic models suggest that the mutation rate would have to be high (so that, there was a sufficient probability of a future knockout of one copy) and effective population size would have to be large (to make selection effective compared to drift; see Elena *et al*., 2007). One study suggested that the product of the mutation rate and the effective population size would have to be >30 for this to be an effective mechanism, several orders of magnitude higher than in the values that have been estimated for mammals (Forster *et al*., 2006). In the absence of such selection, duplicated genes are expected to (and generally observed to) diverge rapidly (Lynch and Conery, 2003; Hughes and Liberles, 2007). So, given the period since eutherian mammalian radiation as time for the duplicates to diverge, is the complementary behavior of selection on the *TRIM22* and *TRIM5* genes due to residual redundancy?

The authors speculate on several mechanisms for the complementarity. One possibility is that the two proteins have evolved different specificities and the complementarity reflects the action of different viruses on different lineages. The authors rightly downplay this hypothesis, regarding it as too simplistic, where nonoverlapping specificities correspond to nonoverlapping viral threat. Instead they propose an alternative based on linkage of the two genes coupled to Hill–Robertson effects of their mutations. Under this hypothesis, selection on the mutation with the strongest adaptive effect will cause not only its fixation, but also the fixation of other variants that had the good fortune to be linked to it, most often because they were present on the haplotype where this advantageous mutation first occurred. These effects on linked sites will extend to the adjacent TRIM gene, and weakens the ability to respond to selection.

An alternative hypothesis draws on the idea of synfunctionalization (see Figure 1; Gitelman, 2007). A canonical example, involves the duplicate TWIST genes that appear to have diverged in function, and subsequently undergone loss of subfunctionalization coupled to gene loss at the regulatory level. A similar evolutionary trajectory has been described in deoxyribonucleoside kinases for changes in protein function (Almgren, 2001; Piskur *et al*., 2004). The substrate specificity of deoxyribonucleoside kinases appears to have undergone subfunctionalization to generate more specific enzymes. However, in the insect lineage, gene copies were lost, coupled with the regeneration of a nonspecific enzyme.

However, if redundancy is lost so quickly for duplicate copies, how can an ancestral function be regained? Some interesting studies of the mutational paths available to proteins evolving new functions, and simulations of protein evolution constrained by protein folding, indicate that the mutational opportunities available to a protein are much more limited than naïve combinatorics would suggest (Bastolla *et al*., 2003; Rastogi *et al*., 2006; Weinreich *et al*., 2006). In the case of the TRIM duplicates, selection is observed in similar locations within the three dimensional structure (within the coiled-coil domain), supporting such an interpreta-

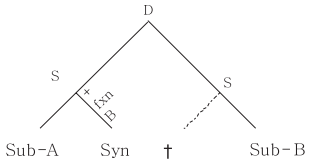

Figure 1 This figure shows the process of synfunctionalization. Here, an ancestral duplication event leads to subfunctionalization. Following a subsequent speciation event, one species retains the subfunctionalized duplicates; the other species loses one copy, whereas the other regains the ancestral function. A less dramatic version of this process can occur at the functional level without gene loss.

tion. If there is residual redundancy coupled to some divergence of function (either at the coding sequence level or as illustrated through nonidentical gene expression patterns), then selection may affect one of the two proteins rather than both. Linkage effects (see above) would further promote synfunctionalization. These possibilities are intriguing and may be the tip of the iceberg comprising links between duplication, divergence, redundancy and the effects of protein structure and function as well as linkage on accessible mutational paths. Whatever the explanation of the anticorrelated positive selection on these TRIM duplicates, the continued characterization of duplicate gene evolution is needed to build on the understanding of gene and proteome evolution that is founded on the seminal Ohno (1970) analysis of gene duplication.

Acknowledgements

I thank Jessica Liberles and Steven Massey for careful reading of this manuscript.

DA Liberles is at the Department of Molecular Biology, University of Wyoming, Laramie, WY 82071, USA.

e-mail: liberles@uwyo.edu

Almgren MAE (2001). *Genomic Approaches to Understanding Protein Function* MSc thesis, Stockholm University, Stockholm, Sweden.

Bastolla U, Porto M, Eduardo Roman MH, Vendruscolo M (2003). Connectivity of neutral networks, overdispersion, and structural conservation in protein evolution. *J Mol Evol* **56**: 243–254.

Cooke J, Nowak MA, Boerlijst M, Maynard-Smith J (1997). Evolutionary origins and maintenance of redundant gene expression during metazoan development. *Trends Genet* **13**: 360–364.

Elena SF, Wilke CO, Ofria C, Lenski RE (2007). Effects of population size and mutation rate on the evolution of mutational robustness. *Evolution* **61**: 666–674.

Forster R, Adami C, Wilke CO (2006). Selection for mutational robustness in finite populations. *J Theor Biol* **243**: 181–190.

Gitelman I (2007). Evolution of the vertebrate twist family and synfunctionalization: a mechanism for differential gene loss through merging of expression domains. *Mol Biol Evol* **24**: 1912–1925.

Hughes AL, Nei M (1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* **335**: 167–170.

Hughes AL, Nei M (1989). Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc Natl Acad Sci USA* **86**: 958–962.

Hughes T, Liberles DA (2007). The pattern of evolution of smaller scale duplicates in mammalian genomes is more consistent with neo- than subfunctionalisation. *J Mol Evol* **65**: 574–588.

Liberles DA, Schreiber DR, Govindarajan S, Chamberlin SG, Benner SA (2001). The adaptive evolution database (TAED). *Genome Biol* **2**, research0028.1–0028.6.

Lynch M, Conery JS (2003). The origins of genome complexity. *Science* **302**: 1401–1404.

Nisole S, Stoye JP, Saib A (2005). TRIM family proteins: retroviral restriction and antiviral defence. *Nat Rev Microbiol* **3**: 799–808.

Ohno S (1970). *Evolution by Gene Duplication*. Springer-Verlag: New York.

Piskur J, Sandrini MP, Knecht W, Munch-Petersen B (2004). Animal deoxyribonucleoside kinases: 'forward' and 'retrograde' evolution of their substrate specificity. *FEBS Lett* **560**: 3–6.

Rastogi S, Reuter N, Liberles DA (2006). Evaluation of models for the evolution of protein sequences and functions under structural constraint. *Biophy Chem* **124**: 134–144.

Sawyer SL, Emerman M, Malik HS (2007). Discordant evolution of the adjacent antiretroviral genes TRIM22 and TRIM5 in mammals. *PLoS Pathog* **3**: e197.

Weinreich DM, Delaney NF, DePristo MA, Hartl DL (2006). Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312**: 111–114.

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 **96**: 39–44.

Van de Peer (2006). 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.