doi:10.1111/j.1558-5646.2011.01412.x

ADAPTIVE *CIS*-REGULATORY CHANGES MAY INVOLVE FEW MUTATIONS

Pablo Razeto-Barry[1,2,3] and Karin Maldonado[1,4]

[1]*Instituto de Filosofía y Ciencias de la Complejidad, IFICC, Los Alerces 3024, Santiago, Chile*

[2]*Universidad Diego Portales, Vicerrectoría Académica, Manuel Rodríguez Sur 415, Santiago, Chile*

[3] *Universidad Andres Bello, Escuela de Ingeniería en Biotecnología, Facultad de Ciencias Biológicas, Fernández Concha 700, Santiago, Chile*

[4]*Departamento de Biología, Facultad de Ciencias, Universidad de La Serena, La Serena, Mailbox 599, Chile*

Received April 17, 2011
Accepted July 7 2011

A long-standing debate in evo–devo research concerns the relative role of protein-coding and *cis*-regulatory regions in adaptation. Recent studies of genetic adaptation have revealed that the number of substitutions contributing to phenotypic variation is lower in *cis*-regulatory than in structural regions, which has led to the idea that *cis*-regulatory regions are less important in phenotypic adaptation. However, the number of substitutions is not the only important factor, the "size" of the adaptive contribution of these substitutions is important too. A geometrical reasoning predicts that, given their lesser pleiotropic effects, *cis*-regulatory substitutions should have a larger average adaptive contribution than protein-coding substitutions. Thus it is possible that even with a lower number of adaptive mutations, *cis*-regulatory regions may contribute at the same level or even more than protein-coding regions.

KEY WORDS: Adaptation, Fitness, Molecular Evolution, Mutations, Pleiotropy, Population Genetics

The Cis-Regulatory Hypothesis

The evolution and development (evo–devo) "*cis*-regulatory hypothesis" claims that adaptive mutations causing phenotypic (especially morphological) evolution are more likely to occur in *cis*-regulatory (noncoding) regions than in protein-coding regions of DNA (see King and Wilson 1975; Carroll 2000, 2005a,b; Wray 2007). There have been several criticisms of this idea, but even the critics recognize that the strongest point in favor of the *cis*-regulatory hypothesis is that mutations in these regions are largely free from pleiotropic (and thus, probably deleterious) effects (Stern 2000; Hoekstra and Coyne 2007; Stern and Orgogozo 2008, 2009).

The consequences of pleiotropy for evolution have been explored theoretically by geometrical frameworks with the Fisher geometrical model of adaptation (Orr 1998, 2000; Welch and Waxman 2003). In this model, the effect of a mutation is represented as a vector in a multidimensional space and pleiotropy of mutations is represented as the number n of orthogonal phenotypic dimensions (Cartesian axes) that the vector can affect (whose magnitude represents the size of its phenotypic effect; Fig. 1). Orr (2000) showed that as the degree of pleiotropy increases, the rate of adaptive evolution decreases dramatically as n^{-1} (see also Welch and Waxman 2003). Orr (2000) showed that universal pleiotropy in complex organisms should greatly reduce the rate of adaptation, calling this effect the "cost of complexity." This cost of complexity is consistent with the *cis*-regulatory hypothesis, given the higher pleiotropy of protein-coding portions of genes compared to *cis*-regulatory regions (Stern and Orgogozo 2008).

The Number of cis-Regulatory Substitutions

In spite of traditional expectations, studies of genetic adaptation have revealed that structural (protein-coding) genes have signals

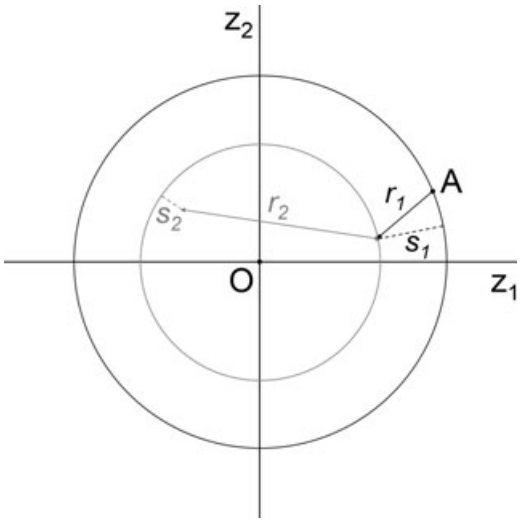

Figure 1. Two-dimensional Fisher's geometrical model contrasting the relationship between mutation size (*r*, the overall phenotypic effect) and selection coefficient (*s*, the adaptive contribution) of substitutions affecting quantitative traits (z_1 and z_2) of a population that has an optimum state in O. The first substitution (dark arrow) has a smaller mutation size (r_1) but larger selection coefficient (s_1) than the second substitution (gray arrow), which has a large mutation size (r_2) with small selection coefficient (s_2).

of a greater number of mutations contributing to phenotypic variation than *cis*-regulatory regions (Hoekstra and Coyne 2007; Stern and Orgogozo 2008). Hoekstra and Coyne (2007) argued that this evidence contrasts with the evo–devo *cis*-regulatory hypothesis, claiming that genetic data lend little support for the presumed predominant role of *cis*-regulatory changes in adaptation (see also Alonso and Wilkins 2005), while Stern and Orgogozo (2008, 2009) restricted this claim to short-term evolution.

Here we argue that the number of mutations alone is not a good measure to compare the adaptive contribution of these two kinds of structures because the average contribution of each mutation may not be the same in both cases. Some time ago King and Wilson (1975) proposed that changes in regulatory regions may be larger than in structural genes such that they could explain major differences between species in spite of the low genetic difference (1%) found between humans and chimpanzees. This argument has been criticized many times, given that in this 1% of differences there are around 60,000 nonsynonymous mutations (Eyre-Walker 2006; Stern and Orgogozo 2008), thus all mutational effects may be of small size and it would not be necessary to invoke regulatory changes. However, beyond this criticism the larger size of regulatory changes by itself is not argument enough for its centrality in phenotypic evolution. In fact, a classical geometrical reasoning predicts that on the average mutations with larger phenotypic effects will probably be more deleterious and more unlikely to be fixed in populations given their lower average adaptive contribution (Fisher 1930; Orr 1998). Consequently, the relevant measure to compare the relative evolutionary contribution of two kinds of mutations is not the average size of the overall phenotypic effect of mutations but the average adaptive contribution of this effect (how much the phenotype advances toward the optimum), typically measured as the selection coefficient of the mutation (Fig. 1).

The Size of cis-Regulatory Adaptive Contributions

Geometrical models have also been used to analyze molecular evolution representing gene pleiotropy as the number n of phenotypic orthogonal axes (complexity) in the Fisher multidimensional space (Gu 2007a,b; Razeto-Barry et al. 2011). Under this framework the equivalent "cost of complexity" implies that on the average an adaptive process involving more pleiotropic genes would occur at a lower rate than one involving less-pleiotropic genes. Surprisingly for us, this geometrical constraint has induced the expectation that "beneficial mutations should be less frequent in complex organisms" (Martin and Lenormand 2006, p. 893), and that "mutations that move the phenotype along only one dimension should contribute to adaptation much more often than mutations that move the phenotype along two or more dimensions simultaneously, that is, mutations that cause pleiotropic effects." (Stern and Orgogozo 2008, p. 2160), while the equivalent conclusion at the protein level can be found in other studies (e.g., Hahn and Kern 2005; Ericson et al. 2006; He and Zhang 2006; Pal et al. 2006; Cooper et al. 2007). That is, several authors have related the lower adaptive rate associated with higher pleiotropy with the expectation of a lower number of substitutions contributing to adaptation. However, we claim that an inversion of reasoning is necessary.

A simple geometrical reasoning shows that the number of adaptive contributions will be "more frequent" for higher pleiotropy (or higher complexity) as $n^{1/2}$ (Box 1, Fig. 2). In fact, Orr (2000) showed that a major source of the lower rate of adaptation for greater complexity is the lower average distance advanced to the optimum by an advantageous substitution. Thus for higher gene pleiotropy a greater number of substitutions is necessary to travel the same distance toward the optimum, and therefore more pleiotropic proteins should spend more time in adaptive processes than less-pleiotropic proteins, each accumulating on the average "a larger number" of mutations with a smaller adaptive contribution (Box 1).

If this protein-centered model applies to regulatory regions, this result suggests that *cis*-regulatory regions should show lower rates of advantageous substitutions than coding regions because they have fewer pleiotropic effects, and not necessarily because

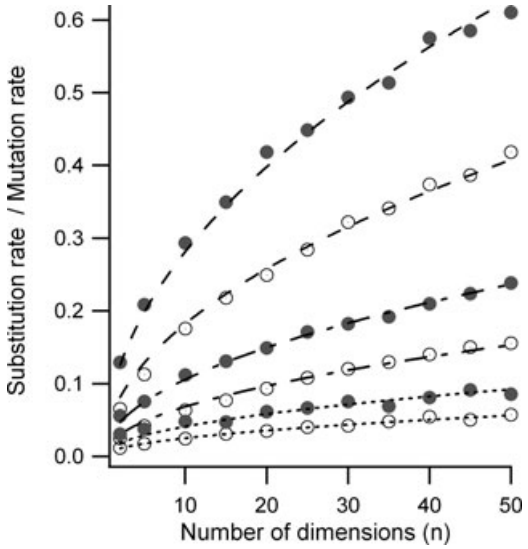

Figure 2. Ratio between advantageous substitution rate and mutation rate (k^+/u) for different numbers of dimensions (n). Open and filled circles correspond to larger ($r = 0.3$) and smaller ($r = 0.1$) average mutations sizes, respectively. From bottom to top the curves correspond to greater variability of random environmental shifts (Poisson process with expected time between changes τ) and their respective curve fittings: dotted lines (for $\tau = 10^6$), for larger and smaller averaged mutation size, $0.0080\sqrt{n}$ and $0.0131\sqrt{n}$, respectively, dot-dashed lines (for $\tau = 3 \cdot 10^5$) $0.0217\sqrt{n}$ and $0.0335\sqrt{n}$, respectively, and dashed lines (for $\tau = 10^5$) $0.0577\sqrt{n}$ and $0.089\sqrt{n}$, respectively. General parameters used were: population size $N = 1000$ and amplitude of environmental variability $\sigma_a = 0.85$. (Modified from Razeto-Barry et al. 2011).

they have a less-significant contribution in phenotypic evolution. If regulatory changes are larger than structural changes (King and Wilson 1975) then this effect would be stronger given that the expected rate of adaptive substitutions is lower for greater mutation size (Fig. 2, Box 1). Thus the relatively large average size (King and Wilson 1975; Tuch et al. 2008) and low pleiotropy of mutations in *cis*-regulatory regions may explain how with fewer mutations these structures can contribute substantially to phenotypic evolution.

It is probable that genetic adaptation involves a combined effect of protein-coding and *cis*-regulatory changes (Hanikenne et al. 2008; Lynch et al. 2008; Tuch et al. 2008). For example, after gene duplication and subfunctionalization, protein-coding mutations could fine-tune some protein function and *cis*-regulatory mutations could specify regulatory expression for this function (Zhang 2003). Thus, we should expect adaptive gene subfunctionalization to involve few *cis*-regulatory and a larger number of protein-coding changes. This prediction could be contrasted with studies such as that of Hanikenne et al. (2008) who showed that adaptation to metal hyperaccumulation in *Arabidopsis halleri* required triplication of a metal pump membrane protein and *cis*-regulatory changes. However, although pairwise comparisons of promoters and coding sequences with its sister species *A. thaliana* were performed (Hanikenne et al. 2008, Table S1) it still remains to be shown which of these substitutions were functional.

Conclusions

We claim that signals of a lower number of advantageous substitutions are not necessarily signals of a lower adaptive contribution. Considering geometrical constraints allows the prediction that average *cis*-regulatory substitutions will have a greater average adaptive contribution than protein-coding nonsynonymous substitutions. Thus in principle *cis*-regulatory regions may be a preponderant cause of phenotypic evolution, because their lower number of adaptive mutations may be balanced by the higher average size of their adaptive contributions. Our theoretical prediction may be tested when selection coefficients for a number of *cis*-regulatory and protein-coding mutations have accumulated. Identifying adaptive mutations is difficult in the first place and the estimation of their selection coefficients is also challenging. Nevertheless, the comparison of selection coefficients calculated for *cis*-regulatory versus structural mutations may be possible using different methods and across different evolutionary timescales. For example, fitness effects on individual genes with new mutations are beginning to be available for measurements in natural species by next generation sequencing (Stapley et al. 2010). On the other hand, using sequence data it has been possible to estimate selection coefficients for particular classes of mutations at the genomic scale (Nielsen 2005). Thus, understanding differences in the size of phenotypic effect and adaptive contribution of different kinds of substitutions may renew the idea that major adaptive shifts can take place by molecular changes without accelerated protein or DNA evolution (King and Wilson 1975), and also may help to explain why the rate of molecular evolution seems to be decoupled from the rate of phenotypic evolution (e.g., Meyer et al. 1990; Sturmbauer and Meyer 1992). These seem to be necessary steps to understand the relationship between genetic evolution and phenotypic evolution, which in turn is necessary to make evolution more predictable and explanatory (Stern and Orgogozo 2008, 2009; Razeto-Barry and Frick 2011).

ACKNOWLEDGMENTS

We thank D. Cotoras for his commentaries on an earlier version of the article.

Box 1. The rate of adaptive substitutions increases with pleiotropy

Let D be the Cartesian distance between the wild-type (A) and the optimum phenotype (O) (Fig. 1). If time is measured in the

scale of the rate of mutations, then the ratio between the number of advantageous mutations fixed in a population (k^+) and the number of mutations (μ) during an adaptive bout (k^+/μ) will be equal to D/d, where d is the average distance that phenotype advances toward the optimum when an advantageous mutation is fixed. Orr (2000) showed that $d = (\sqrt{\pi/2})(r/\sqrt{n})$, where r is the size of mutations. Therefore, the rate of advantageous substitutions in an adaptive bout should be $k^+/\mu = (D/r)(\sqrt{2/\pi})\sqrt{n}$.

Simulations of molecular evolution in which the optimum is randomly moved by environmental changes show a selective scenario in which the majority of substitutions are fixed by positive natural selection and the rate of advantageous substitutions increases with gene pleiotropy as $n^{1/2}$, being this effect stronger for smaller mutation sizes (Fig. 2, see Razeto-Barry et al. 2011).

LITERATURE CITED

Alonso, C. R., and A. S. Wilkins. 2005. The molecular elements that underlie developmental evolution. Nat. Rev. Genet. 6:709–715.

Carroll, S. B. 2000. Endless forms: the evolution of gene regulation and morphological diversity. Cell 101:577–580.

———. 2005a. Evolution at two levels: on genes and form. PLoS Biol. 3:1159–1166.

———. 2005b. Endless forms most beautiful: the new science of evo devo. W. W. Norton & Co., New York.

Cooper, T. F., E. A. Ostrowski, and M. Travisano. 2007. A negative relationship between mutation pleiotropy and fitness effect in yeast. Evolution 61:1495–1499.

Ericson, E., I. Pylvänäinen, L. Fernandez-Ricaud, O. Nerman, J. Warringer, and A. Blomberg. 2006. Genetic pleiotropy in *Saccharomyces cervisiae* quantified by high-resolution phenotypic profiling. Mol. Gen. Genomics. 275:605–614.

Eyre-Walker, A. 2006. The genomic rate of adaptive evolution. Trends Ecol. Evol. 21:569–575.

Fisher R. A. 1930. The genetical theory of natural selection, Oxford Univ. Press, Oxford.

Gu, X. 2007a. Stabilizing selection of protein function and distribution of selection coefficient among sites. Genetica 130:93–97.

———. 2007b. Evolutionary framework for protein sequence evolution and gene pleiotropy. Genetics 175:1813–1822.

Hahn, M. W., and A. D. Kern. 2005. Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks. Mol. Biol. Evol. 22:803–806.

Hanikenne, M., I. N. Talke, M. J. Haydon, C. Lanz, A. Nolte, P. Motte, J. Kroymann, D. Weilgel, and U. Kramer. 2008. Evolution of metal hyperaccumulation required *cis*-regulatory changes and triplication of HMA4. Nature 453:391–396.

He, X., and J. Zhang. 2006. Toward a molecular understanding of pleiotropy. Genetics 173:1885–1891.

Hoekstra, H. E., and J. A. Coyne. 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution 61:995–1016.

King M. C., and A. C. Wilson. 1975. Evolution at two levels in humans and chimpanzees. Science 188:107–116.

Lynch, V. J., A. Tanzer, Y. Wang, F. C. Leung, B. Gellersen, D. Emera, and G. P. Wagner. 2008. Adaptive changes in the transcription factor HoxA-11 are essential for the evolution of pregnancy in mammals. Proc. Natl. Acad. Sci. USA 105:14928–14933.

Martin, G., and T. Lenormand. 2006. A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. Evolution 60:893–907.

Meyer, A., T. D. Kocher, P. Basasibwaki, and A. C. Wilson. 1990. Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. Nature 347:550–553.

Nielsen, R. 2005. Molecular signatures of natural selection. Annu. Rev. Genet. 39:197–218.

Orr, A. H. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52:935–949.

———. 2000. Adaptation and the cost of complexity. Evolution 54: 13–20.

Pal, C., B. Papp, and M. J. Lercher. 2006. An integrated view of protein evolution. Nat. Rev. Genet. 7:337–348.

Razeto-Barry, P., J. Díaz, D. Cotoras, and R. A. Vásquez. 2011. Molecular evolution, mutation size and gene pleiotropy: a geometric reexamination. Genetics 187:877–885.

Razeto-Barry, P., and R. Frick. 2011. Probabilistic causation and the explanatory role of natural selection. Stud. Hist. Phil. Biol. Biomed. Sci. 42:344–355.

Stapley, J., J. Reger, P. G. D. Feulner, C. Smadja, J. Galindo, R. Ekblom, C. Bennison, A. D. Ball, A. P. Beckerman, and J. Slate. 2010. Adaptation genomics: the next generation. Trends Ecol. Evol. 25:705–712.

Stern, D. L. 2000. Evolutionary developmental biology and the problem of variation. Evolution 54:1079–1091.

Stern, D. L., and V. Orgogozo. 2008. The loci of evolution: how predictable is genetic evolution? Evolution 62:2155–2177.

———. 2009. Is genetic evolution predictable? Science 323:746–751.

Sturmbauer, C., and A. Meyer. 1992. Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. Nature 358: 578–581.

Tuch, B. B., H. Li, and A. D. Johnson. 2008. Evolution of eukaryotic transcription circuits. Science 319:1797–1799.

Welch, J. J., and D. Waxman. 2003. Modularity and the cost of complexity. Evolution 57:1723–1734.

Wray, G. A. 2007. The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. 8:206–216.

Zhang, J. 2003. Evolution by gene duplication: an update. Trends Ecol. Evol. 18:292–298.

Associate Editor: C. Alex Buerkle