JOURNAL OF EXPERIMENTAL ZOOLOGY (MOL DEV EVOL) 308B:322–324 (2007)

Energy Costs Constrain the Evolution of Gene Expression

ANDREAS WAGNER*
Department of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland

ABSTRACT Changes in gene expression affect the energy budget of a cell. A previous contribution estimated the energetic cost of yeast mRNA and protein expression from very limited information on protein half-lives. Using recently published genome-scale measurements of protein half-lives, I here confirm that even small increases in gene expression are opposed by natural selection. In small organisms with large effective population sizes, the evolution of transcription and translation rates are thus not evolutionarily neutral processes. *J. Exp. Zool. (Mol. Dev. Evol.) 308B:322–324, 2007.* © 2007 Wiley-Liss, Inc.

How to cite this article: Wagner A. 2007. Energy costs constrain the evolution of gene expression. J. Exp. Zool. (Mol. Dev. Evol.) 308B:322–324.

The syntheses of nucleotides and amino acids, as well as their polymerization into RNA and proteins cost energy. Changes in gene expression thus affect a cell's energy budget. Increases in gene expression can be caused by mutations in regulatory DNA or by gene duplications. These genetic changes can cause substantial changes in gene expression on short evolutionary time scales (Oleksiak et al., 2002; Townsend et al., 2003; Fay et al., 2004; Wittkopp et al., 2004). Is the energy cost of such changes negligible or significant? In other words, are such changes selectively neutral? In a previous contribution, I showed that even the two-fold gene expression changes caused by gene duplication are not neutral (Wagner, 2005). For the median yeast gene, a greater than 10% increase in RNA or protein expression is visible by natural selection (Wagner, 2005). To arrive at these estimates, I had used genome-scale data on the biosynthetic cost of amino acids and nucleotides, mRNA expression, mRNA half-lives, ribosome occupancy of mRNAs, protein abundance, and yeast biomass composition to estimate the energy cost s of gene expression in units of activated phosphate bonds ($\sim$P) as a fraction of a cell's energy budget (Wang et al., 2002; Arava et al., 2003; Forster et al., 2003; Ghaemmaghami et al., 2003; Huh et al., 2003; Hurowitz and Brown, 2004). I estimated the ''critical'' selection coefficient s of mutations that are more strongly influenced by natural selection than by genetic drift from data on synonymous nucleotide polymorphisms and on the mutation rate (Wagner, 2005).

The major limitation of the previous analysis was woefully rudimentary information on protein half-lives. Protein half-lives can vary by more than three orders of magnitude. Because more energy goes into protein expression than into mRNA expression, such variation could seriously affect the energy cost distribution of gene expression. Very recently, genome-scale measurements of protein half-lives have become available (Belle et al., 2006), which I use in the present analysis. Based on these new data, Figure 1a shows the distribution of the energy cost as a fraction of the cell's total energy budget associated with the simultaneous doubling of mRNA and protein expression of a single gene, as might occur after a gene duplication event. The data show that gene duplication incurs an energy cost visible to natural selection for all yeast genes (Wagner, 2005). Figure 1b shows the distribution of the amount of change in protein expression that is neutral. Similar to the earlier study, this value is very

*Correspondence to: A. Wagner, Department of Biochemistry, University of Zurich, Building Y27, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. E-mail: aw@bioc.unizh.ch

Received 6 October 2006; Revised 30 November 2006; Accepted 3 December 2006
Published online 14 March 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jez.b.21152.

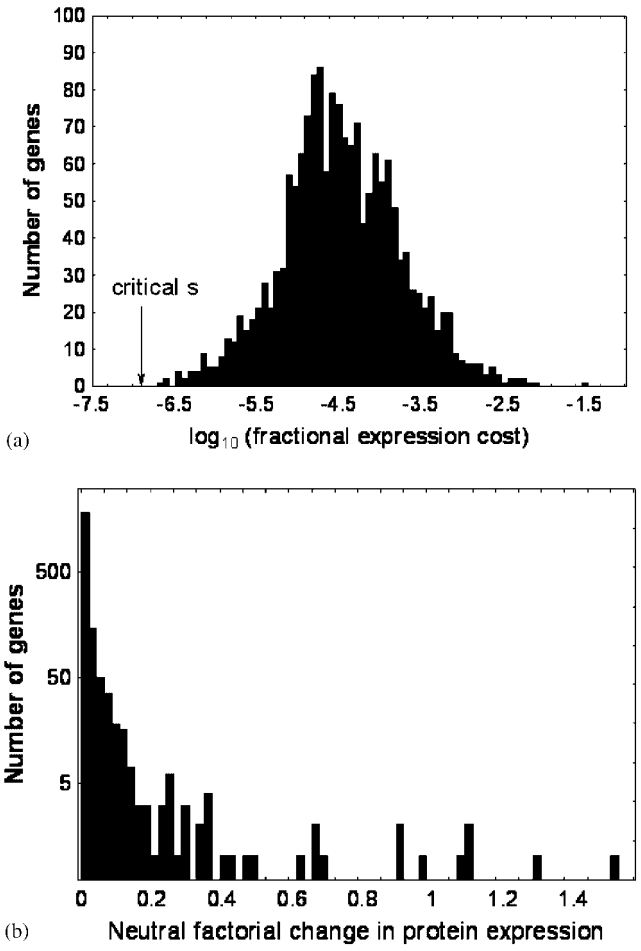

Fig. 1. (**a**) Distribution of the fractional energy cost of simultaneously doubling mRNA and protein expression of a gene. Note the logarithmic scale on the horizontal axis. The arrow points to the fractional cost s below which a change is effectively neutral. Calculation of the critical selection coefficient $s = 1.46 \times 10^{-7}$ below which energy cost changes are neutral relies on the relation $v = 4\mu/\pi$ for haploid organisms (Hedrick, 2000), where $\pi = 0.003$ is the synonymous nucleotide diversity from *S. paradoxus*, *S. cerevisae*'s closest wild relative and $\mu = 2.2 \times 10^{-10}$ the mutation rate (Drake et al., 1998; Johnson et al., 2004), and takes into account that RNA and protein synthesis may account for approximately half of a cell's energy budget (Wagner, 2005). No data on *synonymous* nucleotide diversity are available for *S. cerevisiae* itself, but a recent estimate (Aa et al., 2006) on *overall* nucleotide diversity of $\pi = 0.0046$ (which is typically smaller than synonymous diversity) suggests an upper bound of $s = 9.55 \times 10^{-8}$, rendering the critical s I use here conservative. I used estimates of amino acid biosynthetic costs for respiratory and fermentative conditions from Wagner (2005), except for lysine, where I now take into account that yeast uses α-ketoglutarate instead of oxaloacetate as the lysine precursor. Costs are shown for respiratory conditions, but fermentative conditions yield similar results. (**b**) Distribution of fractional changes in protein expression that are effectively neutral.

small, with a median and maximum of 4.8×10^{-3} and 1.53, respectively. This means that the average (median) yeast protein can change its expression only by 0.5% without a change in energy costs visible to natural selection. Because less energy is invested into mRNA expression, the amount of neutral change that can be tolerated is higher but still very small (median/maximum: 0.035/1.59).

In sum, new genome-scale data confirm that both mRNA and protein expression can change neutrally by only small amounts in yeast. Significant gene expression differences found in comparisons of microbial species with large effective population sizes therefore are influenced by natural selection (Townsend et al., 2003; Fay et al., 2004). More generally, this will hold for all organisms where effective population sizes are large, and where rapid reproduction is coupled to an efficient energy metabolism.

LITERATURE CITED

Aa E, Townsend J, Adams R, Nielsen K, Taylor J. 2006. Population structure and gene evolution in *Saccharomyces cerevisiae*. Fems Yeast Res 6:702–715.

Arava Y, Wang YL, Storey JD, Liu CL, Brown PO, Herschlag D. 2003. Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 100:3889–3894.

Belle A, Tanay A, Bitincka L, Shamir R, O'Shea E. 2006. Quantification of protein half-lives in the budding yeast proteome. Proc Natl Acad Sci USA 103:13004–13009.

Drake JW, Charlesworth B, Charlesworth D, Crow JF. 1998. Rates of spontaneous mutation. Genetics 148:1667–1686.

Fay JC, McCullough HL, Sniegowski PD, Eisen MB. 2004. Population genetic variation in gene expression is associated with phenotypic variation in *Saccharomyces cerevisiae*. Genome Biol 5:R26.

Forster J, Famili I, Fu P, Palsson B, Nielsen J. 2003. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. Genome Res 13:244–253.

Ghaemmaghami S, Huh W, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. 2003. Global analysis of protein expression in yeast. Nature 425:737–741.

Hedrick P. 2000. Genetics of populations. Sudbury, MA: Jones and Bartlett.

Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. 2003. Global analysis of protein localization in budding yeast. Nature 425:686–691.

Hurowitz EH, Brown PO. 2004. Genome-wide analysis of mRNA lengths in *Saccharomyces cerevisiae*. Genome Biol 5:R2.

Johnson LJ, Koufopanou V, Goddard MR, Hetherington R, Schafer SM, Burt A. 2004. Population genetics of the wild yeast *Saccharomyces paradoxus*. Genetics 166:43–52.

Oleksiak MF, Churchill GA, Crawford DL. 2002. Variation in gene expression within and among natural populations. Nat Genet 32:261–266.

Townsend JP, Cavalieri D, Hartl DL. 2003. Population genetic variation in genome-wide gene expression. Molec Biol Evol 20:955–963.

Wagner A. 2005. Energy constraints on the evolution of gene expression. Molec Biol Evol 22:1365–1374.

Wang YL, Liu CL, Storey JD, Tibshirani RJ, Herschlag D, Brown PO. 2002. Precision and functional specificity in mRNA decay. Proc Natl Acad Sci USA 99:5860–5865.

Wittkopp PJ, Haerum BK, Clark AG. 2004. Evolutionary changes in cis and trans gene regulation. Nature 430:85–88.