

Part I. Variation in evolutionary rates across the genome

Spontaneous point mutations arise from errors in DNA replication (that are left uncorrected) and occasionally on exposure to mutagenic chemicals and radiation (Maki 2002). Were these the only factors that dictate the rate of point mutation, one might expect the mutation rate to be uniform across the genome.

To test whether this is true, one assays the rate of nucleotide substitution at sites that are presumed to evolve neutrally. There are advantages and disadvantages to using any class of site (Ellegren, Smith & Webster 2003). In mammals, it has long been assumed that synonymous sites are neutral (Sharp et al. 1995; Duret 2002) and are thus good candidates for representing the mutation rate (Kumar & Subramanian 2002).

Kumar and Subramanian (2002) argued that reports (Wolfe, Sharp & Li 1989; Matassi, Sharp & Gautier 1999; Lercher, Williams & Hurst 2001) that the underlying mutation rate differs among genomic regions (Filipski 1988; Ellegren, Smith & Webster 2003) were unreliable, for several reasons. Most notably, they suggested that regional variation along autosomes (and its corollary, local similarity) is an artefact of the analysis of gene pairs where there is “disparity” in the substitution patterns experienced by each orthologue. Such patterns can result from a gene moving to a new chromosomal region, whereby there may be “amelioration” of its nucleotide content to that of its new genomic location. Consequently, the authors suggested that all genes experience the same underlying mutation rate, that synonymous sites evolve neutrally and that all between-gene variation in evolutionary rate is attributable to estimation error owing to differences in length of sequence.

In Chapter 2, I first ask whether local similarity in rates of evolution is owing to the inclusion of disparate gene pairs. Using both four-fold sites and introns, I find that this is not the case. Even allowing for disparity, neighbouring genes have more similar synonymous rates of evolution than expected by chance. Introns from the same gene also show correlated evolution, including in the rate of insertions/deletions (indels), which should not be susceptible to disparity. That introns from the same gene evolve at similar rates is consistent with investigations of regional variation derived from other forms of non-coding DNA (Chen & Li 2001; Ebersberger et al. 2002; Smith, Webster & Ellegren 2002; Waterston et al. 2002; Hardison et al. 2003; Malcom, Wyckoff & Lahn 2003). The discrepancy with the results of Kumar and Subramanian (2002) appear to be due to their use of a method that is not powerful enough to detect the weak effects of local similarity.

Following the rejection of the notion that local similarity is an artefact, I then go on to investigate the cause of the effect. It does not appear to be due to the observed clustering of genes expressed in the germline, but at least for the rate of synonymous evolution, may be due to the mutagenic effects of meiotic recombination (Perry & Ashworth 1999; Lercher & Hurst 2002; Filatov & Gerrard 2003; Hellmann et al. 2003).

Note that autosomes and sex chromosomes spend different amounts of time in the male and female germlines (Shimmin, Chang & Li 1993) and hence are affected by different mutational forces (e.g. McVean & Hurst 1997; Ebersberger et al. 2002). Consequently, Chapter 2 is only concerned with analysing autosomes.

References

- Chen, F. C., & Li, W. H. (2001) Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *Am. J. Hum. Genet.* **68**: 444-456.
- Duret, L. (2002) Evolution of synonymous codon usage in metazoans. *Curr. Opin. Genet. Dev.* **12**: 640-649.
- Ebersberger, I., Metzler, D., Schwarz, C. & Paabo, S. (2002) Genomewide comparison of DNA sequences between humans and chimpanzees. *Am. J. Hum. Genet.* **70**: 1490-1497.
- Ellegren, H., Smith, N. G. & Webster, M. T. (2003) Mutation rate variation in the mammalian genome. *Curr. Opin. Genet. Dev.* **13**: 562-568.
- Filatov, D. A., & Gerrard, D. T. (2003) High Mutation rates in human and ape pseudoautosomal genes. *Gene* **317**: 67-77.
- Filipski, J. (1988) Why the rate of silent codon substitution is variable within a vertebrates's genome. *J. Theor. Biol.* **134**: 159-164.
- Hardison, R. C. et al. (2003) Covariation in frequencies of substitution, deletion, transposition, and recombination during eutherian evolution. *Genome Res.* **13**: 13-26.
- Hellmann, I., Ebersberger, I., Ptak, S. E., Paabo, S. & Przeworski, M. (2003) A neutral explanation for the correlation of diversity with recombination rates in humans. *Am. J. Hum. Genet.* **72**: 1527-1535.
- Kumar, S., & Subramanian, S. (2002) Mutation rates in mammalian genomes. *Proc. Natl Acad. Sci. USA* **99**: 803-808.
- Lercher, M. J., Williams, E. J. B. & Hurst, L. D. (2001) Local similarity in evolutionary rates extends over whole chromosomes in human-rodent and mouse-rat comparisons: Implications for understanding the mechanistic basis of the male mutation bias. *Mol. Biol. Evol.* **18**: 2032-2039.

- Lercher, M. J., & Hurst, L. D. (2002) Human SNP variability and mutation rate are higher in regions of high recombination. *Trends Genet.* **18**: 337-340.
- Maki, H. (2002) Origins of spontaneous mutations: specificity and directionality of base-substitution, frameshift, and sequence-substitution mutageneses. *Annual Reviews of Genetics* **36**: 279-303.
- Malcom, C. M., Wyckoff, G. J. & Lahn, B. T. (2003) Genic mutation rates in mammals: local similarity, chromosomal heterogeneity, and X-versus-autosome disparity. *Mol. Biol. Evol.* **20**: 1633-1641.
- Matassi, G., Sharp, P. M. & Gautier, C. (1999) Chromosomal location effects on gene sequence evolution in mammals. *Curr. Biol.* **9**: 786-791.
- McVean, G. T., & Hurst, L. D. (1997) Evidence for a selectively favourable reduction in the mutation rate of the X chromosome. *Nature* **386**: 388-392.
- Perry, J., & Ashworth, A. (1999) Evolutionary rate of a gene affected by chromosomal position. *Curr. Biol.* **9**: 987-989.
- Sharp, P. M., Averof, M., Lloyd, A. T., Matassi, G. & Peden, J. F. (1995) DNA-sequence evolution: the sounds of silence. *Phil. Trans. R. Soc. Lond. B* **349**: 241-247.
- Shimmin, L. C., Chang, B. H. & Li, W.-H. (1993) Male-driven evolution of DNA sequences. *Nature* **362**: 745-747.
- Smith, N. G. C., Webster, M. T. & Ellegren, H. (2002) Deterministic mutation rate variation in the human genome. *Genome Res.* **12**: 1350-1356.
- Waterston, R. H. et al. (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**: 520-562.
- Wolfe, K. H., Sharp, P. M. & Li, W. H. (1989) Mutation rates differ among regions of the mammalian genome. *Nature* **337**: 283-285.