Current Biology, Vol. 14, R700–R701, September 7, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.cub.2004.08.041

Natural Selection: Sign, Sign, Everywhere a Sign

Dispatch

Stephen Wooding

Natural selection is an important factor influencing variation in the human genome, but most genetic studies of natural selection have focused on variants with unknown phenotypic associations. This trend is changing. New studies are rapidly revealing the effects of natural selection on genetic variants of known or likely functional importance.

Throughout their evolutionary history, human populations have been exposed to myriad pressures from natural selection. Differential reproductive success arising from variable metabolic efficiency, resistance to pathogens, and even the preferences of potential mates has been brought to bear on our genes. Our genes bear the scars of this pressure — signs that natural selection has been at work — in the form of varying patterns of diversity [1]. Selective sweeps, for instance, which drive favored variants to fixation, tend to reduce diversity near the selected variant and increase the extent of linkage disequilibrium. In contrast, balancing natural selection, which maintains alternative variants at intermediate frequencies, tends to increase diversity near the variant under pressure.

Efforts to read these signs are nothing new. Indeed, some early studies, such as analyses of variation in β-globin and MHC loci, have become classics [2,3]. However, increases in our understanding of the relationship between gene structure and function over the last few years have led to a marked upswing in the number of studies which have investigated the evolutionary history of variants with established phenotypic — and often clinically important — effects. These variants are particularly interesting from an evolutionary standpoint because they are where the phenotypic rubber meets the road of natural selection — variants upon which natural selection could be having particularly direct effects. Studies of such variants pave an unusually direct path between ancient human history and modern human health. The newest addition to this exciting trend appears in this issue of *Current Biology*, in which Rockman *et al.* [4] describe the evolutionary history of a variant in the 5′ regulatory region of the stromelysin gene, *MMP3*.

MMP3 encodes a matrix metalloproteinase that participates in the development and maintenance of the vascular system, including arterial walls, where it has effects on elasticity and thickness. Several studies have found that variance in these traits is associated with a variant composed of a series of either five (5T) or six (6T) thymidine residues in the 5′ regulatory region of *MMP3*. Further, this variant has been found to have significant associations with the development of atherosclerotic plaques, a serious health problem, with the 5T allele being correlated with increased susceptibility to plaque formation but decreased susceptibility to myocardial infarction and aneurism. Given the potentially serious health consequences of this variant, Rockman *et al.* [4] wanted to know whether signs of natural selection occur near the 5T/6T variant in *MMP3*.

To find out, Rockman *et al.* [4] examined geographical variation in 5T/6T allele frequencies, to determine whether human populations are as similar with respect to variation at the *MMP3* locus as they are with respect to a set of control small non-coding polymorphisms (SNPs), and whether any *MMP3* alleles have changed rapidly in frequency, as might be expected during a selective sweep. Analyses of F_{ST}, a measure of population substructuring, revealed that human populations differ much more with respect to variation in *MMP3* than they do with respect to control SNPs, and that Europeans have significantly elevated frequencies of the high-expression 5T allele. In addition, the 5T allele in Europeans is found on a haplotype background that is highly homogeneous relative to that of the 6T allele, a second pattern consistent with a rapid increase in the frequency of 5T. Rockman *et al.* [4] argue that these patterns suggest natural selection has recently driven 5T to high frequencies in Europe — which accounts for the homogeneous haplotype background of 5T — but not other regions of the world — which accounts for the unusually large differences among populations with respect to *MMP3*.

Signs of selection around functionally important variants, such as the 5T/6T *MMP3* alleles, are intrinsically interesting and also important, for they guide us toward new questions and hypotheses of biomedical relevance. Given that the 5T *MMP3* allele has been favored in Europe, for instance, one might hypothesize that 5T provides relative advantages only in temperate environments. Thus, one would like to know: has the 5T allele risen in frequency in other temperate environments, such as in northern North America? Is the 5T allele associated with different phenotypes in different environments?

The results of Rockman *et al.* [4] are but one example of a growing number of studies which focus on specific, functionally important, variants in human genes. Nakajima *et al.* [5], for example, investigated another gene involved in cardiovascular disease, *AGT*. *AGT* encodes angiotensinogen, which participates in blood-pressure homeostasis through its effects on salt retention. At the *AGT* locus, an A(–6) promoter variant causes increased gene expression and risk of essential hypertension relative to the alternative G(–6) variant. Using methods similar to those of Rockman *et al.* [4], Nakajima *et al.* [5] showed that the low-expression

Department of Human Genetics, University of Utah, 15 North 2030 East, Salt Lake City, Utah 84112-5330, USA.
E-mail: swooding@genetics.utah.edu

G(–6) variant occurs at higher frequencies in non-African populations than African populations, and that it has reached such high frequencies only recently. These findings are consistent with the longstanding 'salt retention hypothesis', which posits that populations in hot, humid environments — such as tropical Africa — are under pressure to retain sodium chloride, needed to maintain proper electrolyte balance [6]. In this case, one would like to know, have populations in tropical populations outside Africa evolved high frequencies of the A(–6) variant independently? Are these populations also susceptible to essential hypertension?

Similar signs of selection have been found in the vicinity of other variants with major health effects, including susceptibility to hemochromatosis [7], malaria [8–10], and even attention deficit disorders [11,12]. Equally exciting are signs of selection in genes underlying classic traits. For instance, mounting evidence suggests that selective pressures to digest lactose, a milk protein, have favored variants that extend expression of the lactase gene *LCT* into adulthood in cattle-herding populations [13,14], and balancing selection has been reported near the variants in *T2R38* that distinguish the well-known phenylthiocarbamide (PTC) 'taster' and 'non-taster' alleles [15,16]. Evidence for a complex signature of selection in the *OPN1L* 'red' opsin gene, which accounts for variable color perception [17], has also been reported.

That signs of natural selection are found in the genomic vicinity of functionally important variants is not surprising. Because natural selection acts most strongly on those variants that have phenotypic effects, signs of selection should be found most often in regions of the genome that harbor such variants. Of course, this logic is a two-way street. If signs of selection are present, it is argued, a region is likely to harbor variants that have phenotypic effects [1,18–20]. This suggestion raises the intriguing possibility that inferences about natural selection can be used as a tool in linkage and association studies, two key means of localizing the genes that account for variance in specific traits. Linkage and association studies involving candidate genes, for instance, might gain power by focusing on those genes that show signs of selection, which would seem more likely, *a priori*, to harbor functionally important variants. Though still in their early stages, the use of such approaches is clearly not far down the road. All we need to do is follow the signs.

References

1. Bamshad, M., and Wooding, S. (2003). Signatures of natural selection in the human genome. Nat. Rev. Genet. *4*, 99–111.
2. Pagnier, J., Mears, J.G., Dunda-Belkhodja, O., Schaefer-Rego, K.E., Beldjord, C., Nagel, R.L., and Labie, D. (1984). Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa. Proc. Natl. Acad. Sci. USA *81*, 1771–1773.
3. Hughes, A.L., and Nei, M. (1988). Pattern of nucleotide substitution at major histocompatibility complex loci reveals overdominant selection. Nature *335*, 167–170.
4. Rockman, M., Hahn, M.W., Soranzo, N., Loisel, D.A., Goldstein, D.B., and Wray, G.A. Evolutionary genetics of *MMP3* support hypermutation and positive natural selection as shapers of heart disease. Curr. Biol. *14*, this issue.
5. Nakajima, T., Wooding, S., Sakagami, T., Emi, M., Tokunaga, K., Tamiya, G., Ishigami, T., Umemura, S., Munkbhat, B., Jin, F., et al. (2004). Natural selection and population history in the human angiotensinogen gene (*AGT*): 736 complete *AGT* sequences in worldwide chromosomes. Am. J. Hum. Genet. *74*, 898–916.
6. Gleiberman, L. (1973). Blood pressure and dietary salt in human populations. Ecol. Food Nutr. *2*, 143–156.
7. Toomajian, C., Ajioka, R.S., Jorde, L.B., Kushner, J.P., and Kreitman, M. (2003). A method for detecting recent selection in the human genome from allele age estimates. Genetics *165*, 287–297.
8. Tishkoff, S.A., Varkonyi, R., Cahinkinan, N., Abbes, S., Argyopoulos, G., Destro-Bisol, G., Drousiotou, A., Dangerfield, B., Lefranc, G., Loiselet, J., et al. (2001). Haplotype diversity and linkage disequilibrium at human *G6PD*: Recent origin of alleles that confer malarial resistance. Science *293*, 455–462.
9. Sabeti, P.C., Reich, D.E., Higgins, J.M., Richter, D.J., Schaffner, D.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., McDonald, G.J., Ackerman, H.C., et al. (2002). Detecting recent positive selection in the human genome from haplotype structure. Nature *419*, 832–837.
10. Ohashi, J., Naka, I., Patarapotikul, J., Hananantachai, H., Brittenham, G., Looareesuwan, S., Clark, A.G., and Tokunaga, K. (2004). Extended linkage disequilibrium surrounding the hemoglobin e variant due to malarial selection. Am. J. Hum. Genet. *74*, 1198–1208.
11. Ding, Y., Chi, H., Grady, D., Morishima, A., Kidd, J.R., Kidd, K.K., Flodman, P., Spence, M.A., Schuck, S., Swanson, J.M., et al. (2002). Evidence of positive selection acting at the human dopamine receptor D4 locus. Proc. Natl. Acad. Sci. USA *99*, 309–314.
12. Wang, E., Ding, Y.C., Flodman, P., Kidd, J.R., Kidd, K.K., Grady, D.L., Ryder, O.A., Spence, M.A., Swanson, J.M., and Moyzis, R.K. (2004). The genetic architecture of selection at the human dopamine receptor D4 (*DRD4*) gene locus. Am. J. Hum. Genet. *74*, 931–944.
13. Enattah, N.S., Sahi, T., Savilahti, E., Terwilliger, J.D., Peltonen, L., and Järvelä, I. (2002). Identification of a variant associated with adult-type hypolactasia. Nat. Genet. *30*, 233–237.
14. Bersaglieri, T., Sabeti, P.C., Patterson, N., Vanderploeg, T., Schaffner, S.F., Drake, J.A., Rhodes, M., Reich, D.E., and Hirschhorn, J.N. (2004). Genetic signatures of strong recent positive selection at the lactase gene. Am. J. Hum. Genet. *74*, 1111–1120.
15. Kim, U.K., Jorgenson, E., Coon, H., Leppert, M., Risch, N., and Drayna, D. (2003). Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. Science *299*, 1221–1225.
16. Wooding, S., Kim, U., Bamshad, M., Larsen, J., Jorde, L., and Drayna, D. (2004). Natural selection and molecular evolution in *PTC*, a bitter-taste receptor gene. Am. J. Hum. Genet. *74*, 637–646.
17. Verrelli, B.C., and Tishkoff, S.A. (2004). Signatures of selection and gene conversion associated with human color vision variation. Am. J. Hum. Genet., in press.
18. Akey, J.M., Zhang, G., Zhang, K., Jin, L., and Shriver, M.D. (2002). Interrogating a high-density SNP map for signatures of natural selection. Genome Res. *12*, 1805–1814.
19. Kayser, M., Brauer, S., and Stoneking, M. (2003). A genome scan to detect candidate regions influenced by local natural selection in human populations. Mol. Biol. Evol. *23*, 893–900.
20. Beaumont, M.A., and Balding, D.J. (2004). Identifying adaptive genetic divergence among populations from genome scans. Mol. Ecol. *13*, 969–980.