

their data, if generation times are on average shorter for females (ie mothers are generally younger than the babies fathers), which would imply a smaller global effective population size for females compared to males, then greater

female migration or greater male reproductive variability might also be consistent with this pattern ■

Mark D Shriver is at The Pennsylvania State University, University Park, PA, USA.

Evolutionary Genetics

Evolutionary path to the heart

Shu Ye

European Journal of Human Genetics (2005) **13**, 132–133.

doi:10.1038/sj.ejhg.5201330

Published online 10 November 2004

The idea that evolutionary analyses can be a valuable complementary approach for assessing functional importance of genetic variants is not new,^{1–3} but an elegant recent paper in *Current Biology*, which shows that positive selection might have a significant role in shaping population variation in a gene promoter polymorphism implicated in heart disease, illustrates just how useful these studies can be.⁴

Inherited differences in DNA sequences contribute to interindividual variability in anthropometric characteristics, risk of disease, and response to the environment and medication. Such variants occur roughly every 300–500 bp throughout the human genome, which has about 3 billion base pairs in total. However, only a small percentage of the estimated 10 million polymorphisms result in differences in protein amino-acid sequence or in the level of gene expression. Moreover, it is likely that only some of these functional variants are relevant to disease risk and other phenotypic traits. Molecular genetic epidemiological analyses and *in vitro* functional studies are the most common approaches used to assess the significance of particular genetic variants, but this new study illustrates that evolutionary analyses can be similarly effective.

Matthew Rockman and co-workers investigated the evolutionary history of a polymorphism that has been found to be

associated with alternative cardiovascular phenotypes in a number of studies. The polymorphism is located in the promoter region of the matrix metalloproteinase-3 (MMP3) gene, with two alleles identified in humans, one having a run of five thymidines (5T) and the other six thymidines (6T), from nucleotide position-1608 relative to the transcriptional initiation site of the gene. Previous studies have indicated that this poly-T track forms part of two overlapping transcription factor-binding sites. Although the two alleles have similar affinity with the transcription factor ZNF148 (also named ZBP89), which acts as a transcription enhancer, the 5T allele has a lower affinity than the 6T allele for another transcription factor (apparently an NF- κ B P50/P50 dimer) that acts as a transcription repressor.⁵ As a result, MMP3 transcript and protein levels in *ex vivo* tissues are highest in 5T homozygotes, intermediate in heterozygotes, and lowest in 6T homozygotes.⁶

MMP3 (also known as stromelysin) enzymatically degrades various extracellular matrix proteins in the blood vessel wall and elsewhere. So the expression levels of this gene seem to influence the balance between matrix protein synthesis and degradation, which, in turn, affects the amounts of matrix proteins in tissues. Genetic association studies of the MMP3 gene and atherosclerosis (hardening of the arteries) have shown a genotype–phenotype relationship, such that atheromas

References

- 1 Wilder JA *et al*: *Nat. Genet* 2004; **36**: 1122–1125.
- 2 Helgason A *et al*: *Am. J. Hum. Genet.* 2003; **72**: 1370–1388.

(plaques that forms within the walls lining the arteries) in individuals of the 6T/6T genotype tend to be larger, whereas those in individuals of the 5T/6T or 5T/5T genotype are generally smaller but prone to rupture.⁷ Individuals who carry the 5T allele also seem to have greater arterial elasticity and a predisposition to developing coronary artery aneurysm.⁸ These observations are consistent with a model in which there is an imbalance favouring matrix protein degradation in 5T allele carriers, whereas in 6T/6T individuals matrix protein accumulation is favoured.

The authors data⁴ indicate that the poly-T tract mentioned above might lie within a mutational hot spot that has undergone relatively rapid evolution for tens of million years. They compared the human MMP3 gene promoter sequence with those in nine non-human primates, and found that they all contain the poly-T tract but its length differs among the different species. In addition, intraspecies polymorphism was observed in seven of the primate species studied, further supporting that this region is a mutational hot spot.

The study also suggests that natural selection has caused the frequency of the 5T allele in northern Europe to increase. There are considerable differences in 5T allele frequency among human populations: these frequencies range from 0.01 in Cameroon to 0.54 in Sweden. Differences in allele frequency among populations can arise from natural selection and/or genetic drift (change in gene allele frequency due to chance). The former is unique to each locus, whereas the latter affects all autosomal loci equally. To investigate whether the 5T/6T polymorphism has been a target of natural selection, the authors compared the patterns of genetic differentiation at the 5T/6T site among several populations, with reference to patterns at 18 unlinked, neutral polymorphisms that selection is unlikely to affect. The analyses showed

significant differences at the 5T/6T site between England and Cameroon and between England and India: a pattern that indicates that positive selection might have increased the 5T allele frequency in northern Europe. In addition, haplotype analyses showed that there was a significant excess of a common 5T haplotype in a European–American sample but not in an African–American sample: further evidence of positive selection for the 5T allele in Europe.

Of course further studies would be desirable to confirm these findings; however, in themselves these results are certainly indicative that the 5T/6T polymorphism has an important functional effect. As coronary heart disease is a late-onset disorder, it is likely that the cardiovascular phenotypes associated with the 5T/6T polymorphism are pleiotropic effects of natural selection on other conditions that are currently unknown. The MMPs have important roles in many physiological processes in reproduction and development as well as the pathogenesis of a number of diseases.⁹ It is possible that there is an interaction between the *MMP3* gene and specific European environmental factors that influences one or more of these physiological or pathological processes, and so confers a selective advantage.

Similar recent studies have detected signature of natural selection at a number of other polymorphic sites,^{3,10–13} and we

should expect to see many more studies in this area in the near future. Evolutionary analyses like these not only provide insights into the human evolutionary history and the relative contributions of neutral and advantageous mutations to the wealth of polymorphisms in the human genome today, but also help narrow down the sets of genetic variants to be examined in studies that aim to identify causes of disease risk and response to medicine ■

Acknowledgements

We thank Professor Newton Morton for valuable comments

Shu Ye is at the Human Genetics Division, School of Medicine, University of Southampton, Southampton, UK.
E-mail: Shu.Ye@soton.ac.uk

References

- 1 Przeworski M, Hudson RR *et al*: Adjusting the focus on human variation. *Trends Genet* 2000; **16**: 296–302.
- 2 Nielsen R: Statistical tests of selective neutrality in the age of genomics. *Heredity* 2001; **86**: 641–647.
- 3 Akey JM *et al*: Interrogating a high-density SNP map for signatures of natural selection. *Genome Res* 2002; **12**: 1805–1814.
- 4 Rockman MV *et al*: Positive selection on *MMP3* regulation has shaped heart disease risk. *Curr Biol* 2004; **14**: 1531–1539.
- 5 Ye S *et al*: Progression of coronary atherosclerosis is associated with a common genetic variant of the human stromelysin-1 promoter which results in reduced gene expression. *J Biol Chem* 1996; **271**: 13055–13060.
- 6 Medley TL, Kingwell BA, Gatzka CD, Pillay P, Cole TJ: Matrix metalloproteinase-3 genotype contributes to age-related aortic stiffening through modulation of gene and protein expression. *Circ Res* 2003; **92**: 1254–1261.
- 7 Beyzade S *et al*: Influences of matrix metalloproteinase-3 gene variation on extent of coronary atherosclerosis and risk of myocardial infarction. *J Am Coll Cardiol* 2003; **41**: 2130–2137.
- 8 Lamblin N *et al*: Polymorphisms in the promoter regions of *MMP-2*, *MMP-3*, *MMP-9* and *MMP-12* genes as determinants of aneurysmal coronary artery disease. *J Am Coll Cardiol* 2002; **40**: 43–48.
- 9 Sternlicht MD, Werb Z: How matrix metalloproteinases regulate cell behavior. *Annu Rev Cell Dev Biol* 2001; **17**: 463–516.
- 10 Rockman MV *et al*: Positive selection on a human-specific transcription factor binding site regulating *IL4* expression. *Curr Biol* 2003; **13**: 2118–2123.
- 11 Hahn MW *et al*: Population genetic and phylogenetic evidence for positive selection on regulatory mutations at the factor VII locus in humans. *Genetics* 2004; **167**: 867–877.
- 12 Zhou G *et al*: Haplotype structure and evidence for positive selection at the human *IL13* locus. *Mol Biol Evol* 2004; **21**: 29–35.
- 13 Hull J *et al*: Unusual haplotypic structure of *IL8*, a susceptibility locus for a common respiratory virus. *Am J Hum Genet* 2001; **69**: 413–419.

Genomics

Polymorphic landscape of the human genome

Vivian G Cheung

European Journal of Human Genetics (2005) **13**, 133–135.

doi:10.1038/sj.ejhg.5201331

Published online 10 November 2004

A cluster of recent studies show that variation in the number of copies of large-scale insertion and deletion is common in human populations.

In recent years, our understanding of the nature and the extent of genome variation has expanded greatly. With development of techniques to isolate and manipulate DNA, the focus of genetic variation has shifted from protein polymorphism to DNA sequence polymorphism.

The most common form of DNA sequence variation is single nucleotide