

POPULATION GENETICS

SNPs that come in threes

Single-nucleotide polymorphisms (SNPs) are the bread and butter of many studies of sequence variation, and so understanding how they vary is useful to studies of genome evolution and disease susceptibility. Most human SNPs are biallelic — that is, two allelic variants are segregating in the population — but a paper now shows that there are twice as many triallelic SNPs as expected, and puts forward a mutational mechanism by which they might arise.

Hodgkinson and Eyre-Walker looked at the allelic diversity of SNPs in ~900 nuclear genes. When both exons and CpG dinucleotides were excluded (to avoid biases caused, respectively, by natural selection and

a higher mutation rate), the number of triallelic sites in the data set was 113 — twice the number that would be expected if mutations were randomly distributed. There are several possible explanations for this excess, which the authors proceeded to test.

Natural selection on the region of the SNPs is an obvious candidate. Another is that the nature of the nucleotides surrounding a triallelic SNP increases the local mutation rate. But neither of these explanations holds up — biallelic SNPs are not clustered, as one would expect if selection was preventing them from segregating, and the probability of seeing a triallelic SNP was independent of the local sequence context.

If the mutational properties of particular regions are not to blame, then perhaps the explanation lies in the ability of a site, once mutated, to generate a second mutation. This could occur through the resolution of a mismatched heteroduplex during recombination (for example, a G–C

base pair that mutates to an A–C mismatch, and then to an A–G). This is also not the case, however, as there is no correlation between the recombination rate of a gene and the presence of a triallelic SNP.

Although other, untested explanations could account for the observed levels of diversity, the most plausible is a mutational mechanism that leads to the simultaneous creation of two new base pairs at the same site. The mechanism is not known, but simultaneous generation of two mutations might also explain an observed excess of adjacent biallelic SNPs.

This model will be tested in the not-too-distant future by the [1000 Genomes Project](#); the explanation will be borne out if a haplotype analysis of the non-recombining portion of the Y chromosome reveals that two of the alleles at triallelic sites emerged coincidentally.

Tanita Casci



ORIGINAL RESEARCH PAPER Hodgkinson, A.J. & Eyre-Walker, A. Human tri-allelic sites: evidence for a new mutational mechanism? *Genetics* 2 Nov 2009 (doi:10.1534/genetics.109.110510)
FURTHER READING 1000 Genomes Project: <http://www.1000genomes.org>