NEWS & VIEWS

EVOLUTIONARY GENOMICS

A positive becomes a negative

Laurence D. Hurst

Which human genes have been hotspots for positive selection? Analyses of the top candidates reveal, not genes subject to such selection, but genes that have probably been subject to biased DNA repair.

To find out what makes us humans unique, we can look for genes, and ours alone, in which Darwinian positive selection has occurred. In principle, genes thought to be hotspots for positive selection can be discovered by genome scans that pinpoint especially fast evolutionary change in DNA sequences[1–3]. Work by Berglund *et al.*[4] and Galtier *et al.*[5], however, undermines the assumed connection between fast evolution and pervasive positive selection. Instead, it seems that hotspots have probably accelerated evolution by means of a biased DNA repair process, not because the changes were good for us. Indeed, many changes are probably detrimental.

To infer human-specific changes, the two groups compared human genes with those of the chimpanzee and other primates. Every observed difference was a mutation that was initially rare but became common, maybe — but not necessarily — because of positive selection. Standard tests for positive selection assume that mutations become common either through conferring an advantage on the organism (positive selection) or through chance (drift). The two groups used the same methods for finding hotspots of positive selection as were used in previous studies[1–3]. All of these approaches identify DNA sequence for which the rate of evolution is higher than expected, but they differ in how the expected is defined.

The first method identifies sequence that is evolving at superfast rates in humans compared with other primates (Fig. 1). Berglund *et al.*[4] applied two further methods, both of which compare the rate of evolution of a protein with the rate of evolution of sites in the protein's gene where mutations do not affect it (synonymous sites). The first of these methods asks if the ratio of the two rates (K_a/K_s) is unusually high in humans. The second asks if, after controlling for the variation seen between humans, the between-species value of this ratio is unusually high.

Berglund *et al.* report that the three methods typically pinpoint different candidates for positive selection. However, one regularity does appear. Any position in a gene is occupied by one of four nucleotides, A, T, C or G, combinations of which code for amino acids — the building blocks of proteins. Curiously, the top candidates for positive selection show a great excess of nucleotide changes that were ancestrally either A or T but became G or C. Galtier *et al.*[5] find the same effect, and also show that it applies to hotspots in non-human primates.

Why should this be? Positive selection operating on the choice of amino acids should not so consistently prefer a change of AT to GC. Moreover, the bias, although highly localized within genes, is not unique to the protein-coding parts of the gene, but is seen in the intervening noncoding parts as well[4,5]. Both groups conclude that the hotspot genes are not under positive selection at the protein level.

Assuming that some force is driving the transformation of AT to GC, the reasonable inference is that speedy evolutionary change is giving a false signal of positive selection. Both groups also show that a force promoting AT to GC changes can lead to sequences that not only have accelerated rates of change but also have a K_a/K_s ratio that exceeds one, the common acid-test of positive selection. If this were to be the case, it would require that accelerated evolution is especially concentrated at sites in the gene that, if changed, alter the protein. That is especially likely if, ancestrally, the gene's synonymous (non-amino-acid-specifying) sites were GC rich but sites that specified an amino acid were not: the GC biasing force would cause an excess of protein-changing substitutions compared with the

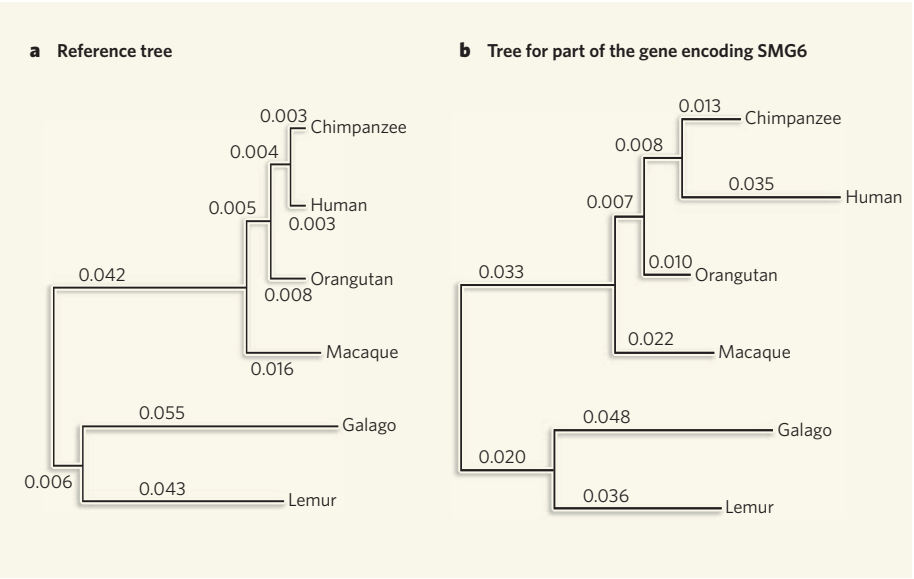

Figure 1 | Identification of a gene segment showing unusually rapid evolution. This example comes from the study by Galtier *et al.*[5], who identified human-specific accelerated change in DNA sequence by comparing the proportional number of changes seen in a given gene segment (or gene) with that of a reference set from humans and other primates. **a**, The reference set drawn from 1,000 genes. Considering that 0.003 changes have occurred in the human-specific lineage, compared with the sum of all branch lengths across the tree (0.185), we conclude that only 1.6% of all sequence change happened after the split between humans and chimps in the human lineage. **b**, Tree resulting from sequence data for part of the gene encoding the SMG6 protein. Here, 15% of all the sequence change has occurred in the human lineage, implying human-specific acceleration of rates. Of 29 changes, 25 are AT→GC, implicating biased gene conversion as the cause. Numbers indicate the average number of amino-acid changes per site. (Figure adapted from ref. 5.)

relatively unaffected synonymous sites, these always being GC rich[4].

What might the biasing force be? Changing nucleotides at synonymous sites modulates expression of a gene[6]. But why then would the changes be highly localized within genes, and why is the bias also in noncoding sequence? A simpler explanation than positive selection on either proteins or expression rate, and one that anticipated[7] the new results, invokes a biased DNA repair process.

During the manufacture of sex cells, a cell with two copies of each of our 23 chromosomes divides to produce cells with just one set of each. During this process, chromosomes can swap DNA (recombination); this involves a break in one chromosome that exposes a single strand of the normally double-stranded DNA. The single strand then finds a complementary strand in its partner chromosome. The two strands pair up to make a new double-stranded bit of DNA. The sequence of the two strands may not, however, be perfectly complementary and might break the rules of DNA pairing (G should pair with C, and A with T). Mismatch repair enzymes then correct rule violations. Imagine a C mismatched with an A. There are two choices for repair: replace C with T or replace A with G. Importantly, the system is biased[8] and more commonly replaces A with G. More generally, it favours Gs and Cs over As or Ts. The repair bias may be an evolved property to cope with a high mutation rate of C to T.

Biased gene conversion[9] (BGC), as the process is termed, explains a general trend towards higher rates of evolution in chromosome domains that commonly undergo recombination[10], and correctly predicts the high rates of recombination in the superfast hotspots[4,5]. Many such sites lie towards the ends of chromosomes, where recombination is common. Moreover, as gene conversion also happens between physically adjacent duplicated genes, BGC explains why chromosomally neighbouring duplicates have similar GC-rich sequence[11]. A suggestive finding is that two top hotspots are a pair of adjacent duplicate genes (those encoding the olfactory receptors OL3A3 and OL3A2).

Importantly, BGC can drive mutations that are deleterious[5]. This may explain why hotspots can occur in genes that, in species other than ourselves, are under strong selection not to change, and hence are not obvious candidates for positive selection. Given that BGC can force deleterious mutations to spread through a population, part of the high rate of evolution in the hotspots could be because of the subsequent spread of compensatory mutations[5].

There remains one mechanistic oddity. It has been observed that the correlation between evolutionary rate and recombination rate holds only for the recombination rate seen in males (see, for example, refs 10 and 12). This is evidence against a role for recombination's randomizing effects (a shuffling of which mutations sit next to each other), which form the underpinnings of a potential alternative explanation. Why the effect is male-specific is not known.

The results of Berglund *et al.*[4] and Galtier *et al.*[5] accord with the view of BGC as a driver of sequence evolution, potentially explaining the occurrence of large spans of approximately homogeneous nucleotide content (isochores) in our genome[13]. More disturbingly, the results bring into question the usefulness of the standard tool kit for identifying hotspots of changes that are beneficial to organisms. Convincing demonstration of positive selection now requires both evidence that the changes were not caused by BGC and scrutiny of the impact of the amino-acid changes. ■

Laurence D. Hurst is in the Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK.
e-mail: l.d.hurst@bath.ac.uk

1. Pollard, K. S. *et al. Nature* **443,** 167–172 (2006).
2. Bustamante, C. D. *et al. Nature* **437,** 1153–1157 (2005).
3. Clark, A. G. *et al. Science* **302,** 1960–1963 (2003).
4. Berglund, J., Pollard, K. S. & Webster, M. T. *PLoS Biol.* **7,** e1000026 (2009).
5. Galtier, N., Duret, L., Glémin, S. & Ranwez, V. *Trends Genet.* **25,** 1–5 (2009).
6. Kudla, G., Lipinski, L., Caffin, F., Helwak, A. & Zylicz, M. *PLoS Biol.* **4,** 933–942 (2006).
7. Galtier, N. & Duret, L. *Trends Genet.* **23,** 273–277 (2007).
8. Brown, T. C. & Jiricny, J. *Cell* **54,** 705–711 (1988).
9. Marais, G. *Trends Genet.* **19,** 330–338 (2003).
10. Duret, L. & Arndt, P. F. *PLoS Genet.* **4,** e1000071 (2008).
11. Galtier, N. *Trends Genet.* **19,** 65–68 (2003).
12. Dreszer, T. R. *et al. Genome Res.* **17,** 1420–1430 (2007).
13. Eyre-Walker, A. & Hurst, L. D. *Nature Rev. Genet.* **2,** 549–555 (2001).

IMMUNOLOGY

Natural killer cells remember

Sophie Ugolini and Eric Vivier

Cells of the adaptive immune system hold a grudge: on re-encountering a pathogen, they show a robust protective response. It seems that natural killer cells of the innate immune system might also have this ability.

Learning, a hallmark of life, produces adaptation to new information. The immune system, like the nervous system, has this ability to learn from previous experience — such as a single encounter with the many pathogens that exist. The result is immunological memory that confers long-lasting protection. For instance, once exposed to the measles virus in childhood, humans are immune to the disease for up to 75 years[1]. Until now, immunological memory was thought to be a feature of the adaptive immune system, specifically, of immune cells called T and B cells. In this issue, however, Sun *et al.*[2] (page 557) shed light on an unexpected player in the persistence of immunity: natural killer cells, which have traditionally been considered to be part of the innate immune system.

Adaptive immunity appeared in vertebrates around 500 million years ago[3]. In the more complex vertebrates, each clone of T and B cells expresses a unique cell-surface receptor, which preferentially recognizes a specific antigen and so, potentially, a specific invading pathogen. The ensemble of antigen-specific T- and B-cell receptors is called the immune repertoire. When these receptors recognize and bind an antigen, the clones of T or B cells on which the receptors are expressed expand dramatically in number, and the cells acquire effector functions — for example, T cells acquire the ability to kill the offending antigen-containing cells and B cells secrete antibodies. Once the immune response is over, a 'contraction' phase occurs, and only a small fraction of the expanded T- or B-cell population survives. These long-lived memory cells are central to maintaining long-term immunity after infection or vaccination. T and B cells are thus said to be adaptive, because their repertoire depends on an individual's antigenic history.

Natural killer cells are classified as part of the innate immune system. Through many receptors, they recognize a vast array of molecules, the expression of which is indicative of a particular situation, such as microbial infection, tumour formation or DNA damage[4]. Sun *et al.*[2] now show that natural killer cells also have properties previously ascribed only to T and B cells.

In mice, a subset of natural killer cells, which express the Ly49H receptor on their surface ($Ly49H^+$ cells), specifically recognize a protein produced by some strains of mouse cytomegalovirus (MCMV)[5,6]. The specific recognition of these cells, an ability also shared by cells of the adaptive immune system[5,6], allows $Ly49H^+$ natural killer cells to limit MCMV infection at an early stage. But before infection, the number of $Ly49H^+$ natural killer cells is much higher than the number of antigen-specific T cells (around half of all natural killer cells are $Ly49H^+$, compared with between 1 in 10^4 and 1 in 10^8 T cells for a given T-cell clone before infection).

Previous work[7] has shown that MCMV infection results in a two- to tenfold increase in the number of natural killer cells. To compare this response under conditions similar to that of a T-cell response, Sun *et al.* experimentally reduced the initial number of $Ly49H^+$ natural