

What makes us human?

Tarjei S Mikkelsen

Address: Broad Institute, Massachusetts Institute of Technology and Harvard University, 320 Charles Street, Cambridge, MA 02139, USA.
E-mail: tarjei@broad.mit.edu

Published: 27 July 2004

Genome Biology 2004, **5**:238

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2004/5/8/238>

© 2004 BioMed Central Ltd

Abstract

The sequence of chimpanzee chromosome 22 is starting to help us to define the set of genetic attributes that are unique to humans, but interpreting the biological consequences of these remains a major challenge.

In a recent paper in *Nature* [1], the International Chimpanzee Chromosome 22 Consortium describes the sequencing and initial analysis of chromosome 22 of the chimpanzee *Pan troglodytes* (PTR22), the ortholog of human chromosome 21 (HSA21, which is involved in the trisomy that leads to Down syndrome). A whole-genome shotgun draft assembly of the entire chimpanzee genome was made public by a US-based consortium in 2003 [2], but the sequence published by Watanabe *et al.* [1] - which was sequenced from bacterial artificial chromosome clones rather than by whole-genome shotgun - represents the first 'finished' chimpanzee chromosome, meaning that its completeness, contiguity and error rates are comparable to the current human genome sequence [3]. In addition to being a valuable quality control for the whole-genome shotgun assembly, finished sequence is better suited for studying insertions, deletions and other structural variation between the human and chimpanzee genomes. The analysis by Watanabe *et al.* [1] also constitutes the first complete and unbiased comparison of a human and a chimpanzee chromosome at the sequence level.

Watanabe *et al.* [1] found that PTR22 and HSA21 differ at approximately 1.44% of their 33 million aligned nucleotides. In addition, they found 68,000 insertions or deletions (indels), the vast majority less than 300 bp in size. The number of indels accumulated over the time since humans and chimpanzees diverged is therefore approximately one seventh the number of point mutations in the same period; as several nucleotides are affected by each indel event, however, this result confirms previous estimates [4] that indels are a

major source of sequence divergence between humans and chimpanzees. Comparison of the human-chimpanzee indels to gorilla and orangutan sequences suggests that both PTR22 and HSA21 have undergone a small net decrease in size since speciation, but it is unclear whether this observation can be extrapolated to the respective complete genomes.

Rapidly evolving proteins

Comparing 231 orthologous genes on the chromosomes, Watanabe *et al.* [1] found 179 cases in which the human and chimpanzee protein-coding sequences were of equal lengths. Of these, approximately 80% have at least one amino-acid difference between the two species, leading to an average amino-acid divergence of 0.82%. Interestingly, of the remaining 52 orthologs, 15 were found to have indels within their coding sequences and 32 were found to have changes in the first ATG (start codon) or the stop codon, changes that would potentially lead to gross structural differences between the human and chimpanzee protein products. Given that fewer than 54% of human-mouse orthologs have coding sequences of different lengths [5], it seems rather surprising that as many as 20% have changed between humans and chimpanzees, despite the significantly shorter time since their divergence. Watanabe *et al.* [1] hypothesized that indels and structural changes may represent one of the major mechanisms of proteome evolution in the higher primates.

But are the data reported by Watanabe *et al.* [1] - based on less than 1% of the known human complement of genes -

representative of human and chimpanzee evolution in general? There are a few observations that suggest caution is needed. First, survey sequencing (sequencing of random short regions over the whole genome) has suggested that HSA21 and PTR22 are diverging faster than most of the other autosomes [6], implying that comparing these chromosomes may overestimate the genome-wide rate of divergence somewhat. Second, approximately two-thirds of the orthologs reported to show length differences between humans and chimpanzees are uncharacterized or poorly characterized genes whose coding sequences have typically been annotated on the basis of a small number of cDNAs or expressed sequence tags. Few of these predicted genes have an unambiguous mouse ortholog, whereas it is estimated that approximately 80% of all human genes do [5]. Genes that differ in length between human and chimp also dominate the list of orthologs that have high ratios of nonsynonymous to synonymous substitutions [1], implying that they are evolving under more relaxed constraints than the average human-chimpanzee ortholog pair.

Thus, a significant proportion of the genes reported to be rapidly diverging by Watanabe *et al.* [1] appear to be novel to the primate lineage and of largely unknown function. In contrast, better characterized genes with known functions and recognizable mouse orthologs are highly conserved. This suggests that the relatively high number of genes with putative structural changes may not be so surprising after all, because earlier estimates of the structural change rate were probably based on highly conserved genes. It also raises the question of the relative contributions to the evolution of higher primates of rapidly changing, 'novel' genes versus genes that are widely conserved in mammals. On the one hand, gene duplication followed by adaptive evolution is one of the major forces for the emergence of new gene functions [7], and a putative example of this phenomenon during the emergence of humans and the African apes has been described previously [8]. On the other hand, the rate and pattern of morphological changes in modern humans and our hominid ancestors may not be all that different from other mammals [9], suggesting that the modification of existing, highly-conserved developmental pathways, rather than the invention of new genes and features, may explain much of human evolution. Further studies are needed to resolve this issue.

Changes in gene expression

It has long been argued that changes in gene regulation may be more important to morphological and functional evolution than overall genomic divergence [10]. Using Affymetrix array technology, Watanabe *et al.* [1] compared the expression profiles of genes on human HSA21 and chimpanzee PTR22 in various tissues and identified 9 and 12 genes with significantly different expression levels between species in brain and liver, respectively. An intriguing example of a

differentially expressed gene is the transcription factor *ETS2*, which is upregulated in chimpanzee brains relative to humans. Subtle upregulation of the orthologous *Ets2* gene in developing mice can lead to cranial and cervical skeletal abnormalities reminiscent of those found in people with Down syndrome [11], but the implications of this finding to human and chimpanzee evolution are unclear.

Interestingly, Watanabe *et al.* [1] also found a correlation between changes in expression levels and changes in sequences upstream of genes. They found that orthologous genes with high divergence in their 5' untranslated regions (UTRs) tend to show differences in expression levels. Similarly, orthologous genes associated with more diverged CpG islands also tend to show different expression levels. These two trends may be related, as CpG islands often overlap with the first exons of genes. It has been proposed that 5' UTRs might be under positive selection in humans, possibly because of their involvement in the regulation of expression levels [12]. In related work, Enard *et al.* [13] recently suggested that DNA methylation patterns, which are important modulators of gene expression and protect CpG dinucleotides from mutation, differ between human and chimpanzee brains. If these differences extend to the germline, they might also explain the correlation between expression and sequence divergence.

But do differences in expression levels of human and chimpanzee genes necessarily have functional consequences? The recent important paper by Khaitovich *et al.* [14] suggests that this may not be the case. Their work shows that gene-expression differences between mammalian species accumulate linearly with time, and that the rate of accumulation does not differ between intact genes and expressed pseudogenes [14]. The implication is that the majority of expression differences observed between two species, like the majority of amino-acid differences, are likely to be selectively neutral or nearly neutral and therefore of little or no functional significance. Thus, the interpretation of species-specific expression differences will need to be based on comparisons with a null model of how expression changes under neutral evolution.

Different perspectives on human evolution

The popular media did not quite know what to make of the initial analysis of PTR22 [1]. "Chimp DNA almost identical to ours" announced Reuters, whereas Asian News International informed its readers that "Chimps and men are indeed very different!" Unwittingly, the authors of these two headlines may have summarized many years' worth of scientific debate over human and chimpanzee evolution [10]. We can now count the exact number of genetic differences between humans and chimpanzees, but whether this number is high or low is entirely in the eye of the beholder. Humans and chimpanzees are an order of magnitude more different, in

terms of genetic changes, than any two humans, but an order of magnitude less different than mice and rats are from each other. And although rat biologists will no doubt disagree, most of us might like to think that what separates us from the chimpanzee is far more profound than what separates a small rodent from a slightly larger rodent.

The major question that is before us now is thus not whether we are as different from other species as we might like to think, but rather which of the human-specific genetic changes account for our unique biological traits and which are simply evolutionary noise. Answering this question will require additional data from other primates as well as fundamental advances in our understanding of the functional evolution of both coding and non-coding sequences. We are a long way from understanding the genetic basis for the origins of the human species, but the sequencing of the chimpanzee genome is an important milestone along the road.

References

1. Watanabe H, Fujiyama A, Hattori M, Taylor TD, Toyoda A, Kuroki Y, Noguchi H, BenKahla A, Lehrach H, Sudbrak R, et al.: **DNA sequence and comparative analysis of chimpanzee chromosome 22.** *Nature* 2004, **429**:382-388.
2. **Chimp genome assembled by sequencing centers** [<http://www.genome.gov/11509418>]
3. Schmutz J, Wheeler J, Grimwood J, Dickson M, Yang J, Caoile C, Bajorek E, Black S, Chan YM, Denys M, et al.: **Quality assessment of the human genome sequence.** *Nature* 2004, **429**:365-368.
4. Britten RJ: **Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels.** *Proc Natl Acad Sci USA* 2002, **99**:13633-13635.
5. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, et al.: **Initial sequencing and comparative analysis of the mouse genome.** *Nature* 2002, **420**:520-562.
6. Ebersberger I, Metzler D, Schwarz C, Pääbo S: **Genomewide comparison of DNA sequences between humans and chimpanzees.** *Am J Hum Genet* 2002, **70**:1490-1497.
7. Ohno S: *Evolution by gene duplication.* Berlin: Springer; 1970.
8. Johnson ME, Viggiano L, Bailey JA, Abdul-Rauf M, Goodwin G, Rocchi M, Eichler EE: **Positive selection of a gene family during the emergence of humans and African apes.** *Nature* 2001, **413**:514-519.
9. Carroll SB: **Genetics and the making of *Homo sapiens*.** *Nature* 2003, **422**:849-857.
10. Cherty LM, Case SM, Wilson AC: **Frog perspective on the morphological difference between humans and chimpanzees.** *Science* 1978, **200**:208-211.
11. Sumarsono SH, Wilson TJ, Tymms MJ, Venter DJ, Corrick CM, Kola R, Lahoud MH, Papas TS, Seth A, Kola I: **Down's syndrome-like skeletal abnormalities in *Ets2* transgenic mice.** *Nature* 1996, **379**:534-537.
12. Hellmann I, Zollner S, Enard W, Ebersberger I, Nickel B, Pääbo S: **Selection on human genes as revealed by comparisons to chimpanzee cDNA.** *Genome Res* 2003, **13**:831-837.
13. Enard W, Fassbender A, Model F, Adorjan P, Pääbo S, Olek A: **Differences in DNA methylation patterns between humans and chimpanzees.** *Curr Biol* 2004, **14**:R148-R149.
14. Khaitovich P, Weiss G, Lachmann M, Hellmann I, Enard W, Muetzel B, Wirkner U, Ansorge W, Pääbo S: **A neutral model of transcriptome evolution.** *PLoS Biol* 2004, **2**:E132.