

# Has the chimpanzee Y chromosome been sequenced?

## To the Editor:

Kuroki *et al.* recently reported “the finished sequence of the chimpanzee Y chromosome”<sup>1</sup>. Their analyses included comparisons with previously reported DNA sequences from the human and chimpanzee Y chromosomes<sup>2,3</sup>. The article<sup>1</sup> was based on the authors’ sequencing of 12.7 Mb from the PTB1 library, which

represents the genome of one male chimpanzee. We previously sequenced the 9.5-Mb ‘X-degenerate’ portion of the Y chromosome from a different male chimpanzee, whose genome is represented in the CHORI-251 library<sup>2</sup>. We write to express concerns regarding the conclusions of Kuroki *et al.*, including the gene content of the chimpanzee and human Y

chromosomes, and the level of sequence divergence between the two chimpanzee Y chromosomes whose sequences have been explored.

First, the authors’ claim of “the finished sequence of the chimpanzee Y chromosome” merits attention<sup>1</sup>. The 12.7 Mb reported in the study overlaps fully the 9.5-Mb X-degenerate region analyzed in the prior study<sup>2</sup>; it also

includes 1.7 Mb of contiguous, non-X-degenerate sequence not examined in the earlier publication. The remaining 1.5 Mb reported by the authors is a superficial sampling of the 'ampliconic' portions of the chimpanzee Y chromosome. The ampliconic regions of primate Y chromosomes are of great biological and medical interest<sup>3–10</sup>. These regions are difficult but not impossible to sequence systematically and comprehensively (see refs. 3 and 5 and our unpublished results), and, in man, they comprise 10.2 Mb, or nearly half of the Y chromosome's male-specific euchromatin<sup>3</sup>. If a similar fraction of the chimpanzee Y chromosome is ampliconic, then large and biologically significant portions of the chromosome have yet to be sequenced and analyzed.

The authors<sup>1</sup> reported more genes within the X-degenerate regions of the chimpanzee and human Y chromosomes than did investigators in earlier studies<sup>2,3</sup>, but these additions, we suggest, do not withstand scrutiny. Unlike prior studies of the human and chimpanzee Y chromosomes<sup>2,3</sup>, the authors' inferences were based on very limited electronic analyses and were not validated experimentally. This may explain why several pseudogenes or disrupted genes—some explicitly identified as such in earlier studies—were treated as functional genes despite previous experimen-

tal evidence to the contrary (**Supplementary Note** and **Supplementary Figure 1** online). These include the *TMSB4Y* and *USP9Y* pseudogenes on the chimpanzee Y chromosome<sup>2</sup>, the *GYG2* pseudogene on the human and chimpanzee Y chromosomes<sup>2,3</sup> and the *CD24LA* pseudogene on the human Y chromosome.

Finally, the authors appear to have overestimated the nucleotide divergence between the two chimpanzee Y chromosomes represented by the PTB1 and CHORI-251 libraries. We aligned the PTB1 and CHORI-251 sequences (**Supplementary Tables 1–3** online; sequence alignments can be found at <http://jura.wi.mit.edu/page>) and found their divergence to be 0.002%, or roughly 20 times lower than the 0.0422% reported by the authors<sup>1</sup>. (The authors similarly overestimated divergence between PTB1 and a third chimpanzee Y chromosome, represented by the RPCI-43 library; our sequence alignments can be found at <http://jura.wi.mit.edu/page>. Note that all CHOR-251 and RPCI-43 sequences included in our alignments with PTB1 were publicly available, as finished sequence, prior to the study by Kuroki *et al.*) Our calculation of divergence between the Y chromosomes of PTB1 and CHORI-251 is so low (~1 in 50,000 nucleotides) that sequencing errors (estimated at

less than 1 in 200,000 nucleotides in each study) could account for about one in every three substitutions that appear to differentiate the chromosomes (**Supplementary Note**). It is unclear how the authors calculated a divergence 20-fold higher than ours when comparing the same sequences.

Jennifer F Hughes<sup>1</sup>, Helen Skaletsky<sup>1</sup>,  
Steve Rozen<sup>1</sup>, Richard K Wilson<sup>2</sup> &  
David C Page<sup>1</sup> <http://jura.wi.mit.edu/rozen>

<sup>1</sup>Howard Hughes Medical Institute, Whitehead Institute, and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, Massachusetts, 02142, USA. <sup>2</sup>Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, Missouri 63108, USA.  
e-mail: [dcpage@wi.mit.edu](mailto:dcpage@wi.mit.edu)

1. Kuroki, Y. *et al. Nat. Genet.* **38**, 158–167 (2006).
2. Hughes, J.F. *et al. Nature* **437**, 100–103 (2005).
3. Skaletsky, H. *et al. Nature* **423**, 825–837 (2003).
4. Yen, P. *Nat. Genet.* **29**, 243–244 (2001).
5. Kuroda-Kawaguchi, T. *et al. Nat. Genet.* **29**, 279–286 (2001).
6. Rozen, S. *et al. Nature* **423**, 873–876 (2003).
7. Hurles, M.E. & Jobling, M.A. *Nat. Genet.* **34**, 246–247 (2003).
8. Anonymous. *Nat. Genet.* **35**, 195 (2003).
9. Tyler-Smith, C. & McVean, G. *Nat. Genet.* **35**, 201–202 (2003).
10. Repping, S. *et al. Nat. Genet.* **35**, 247–251 (2003).

Kuorki *et al.* reply:

We very much appreciate that Dr. Page and his colleagues have spared their valuable time to carefully evaluate and re-analyze the data from our chimpanzee Y chromosome comparative analysis<sup>1</sup>.

First, we would like to clarify several misunderstandings concerning our paper, particularly about the sequenced region in our paper<sup>1</sup>. We determined complete sequences for 271 kb of the Y-specific pseudoautosomal region 1 and 12.7 Mb of the male-specific region. We produced high-quality sequence data for almost half of the entire chimpanzee Y chromosome and carried out detailed comparative analyses between human and chimpanzee Y chromosomes. We contrasted this with similar analyses that were carried out on the autosomes, namely human chromosome 21 and chimpanzee chromosome 22 (now renumbered to 21), and the non-recombining portions of the Y chromosome<sup>1</sup>. In addition to the interspecies analyses, we examined the diversity in chimpanzees using publicly available sequence data<sup>1–3</sup> and verified that the diversity in the chimpanzee Y

chromosome was very low. These analyses include the entire X-degenerate region and parts of the ampliconic region<sup>1</sup>; the remaining parts of the ampliconic region have not yet been fully examined. Our results showed that the structure and sequence identity of the regions were extremely different between human and chimpanzee. We found that both human and chimpanzee Y chromosomes retain the same basic structure: that is, many palindromic structures have accumulated on the Y chromosome, but the regions involved in the palindrome conformation are species specific, such as the chimpanzee-specific palindrome CSP1 (see Supplementary Fig. 6 in ref. 1 and our website, <http://stt.gsc.riken.jp/> (RIKEN Genomic Sciences Center)). We agree it is important to completely sequence the chimpanzee Y chromosome ampliconic region and to carry out more comprehensive comparative analyses.

Concerning the differing results of our gene annotation, this is dependent on the method used for the analysis<sup>1,2</sup>. As described in the papers, we used the human Y chromosome gene set manually annotated by the HAVANA group ([\[vega.sanger.ac.uk/homo\\\_sapiens/\]\(http://vega.sanger.ac.uk/homo\_sapiens/\)\), whereas Hughes \*et al.\* used their own annotation data for human Y chromosome. Hughes \*et al.\* emphasized that our annotation results were wrong and that their results, which were verified by RT-PCR, were more reliable. We feel that Hughes \*et al.\* not only overestimated the reliability of the RT-PCR technology but misunderstood our conclusions: that is, we did not assert whether these four genes, \*USP9Y\*, \*TMSB4Y\*, \*AC002992.5\* \(\*GYG2-like\*\) and \*CD24LA\*, were functional, because it could not be determined from our sequence-based analysis. Even if we had RT-PCR data, that it in itself would not be conclusive as to the functionality of these genes. Further support, such as full-length cDNA cloning or RNA blot analysis, would be more convincing. We agree that further careful analysis for gene identification, annotation and verification is necessary to identify the functional and biological meaning not only of the chimpanzee genes but also of their human counterparts, an area of ongoing research.](http://</a></p>
</div>
<div data-bbox=)

Hughes *et al.* recalculated the chimpanzee diversity and had some concerns about

the difference in our results. We suggest that the difference is due to a combination of alternate calculation methods and slightly different data sets (at the time of our analysis, the assembly of Hughes *et al.*, DP000054, was not available). We used a local alignment method, while they used a global alignment method. The reason we used a local alignment method was to compare the diversity of the Y chromosome with that of the autosomes—in this case, human chromosome 21 and chimpanzee chromosome 22. We re-analyzed the chimpanzee Y chromosome diversity using

different parameters, 90% and 99% identity over a length of 1,000 bp, and obtained the values 0.042% to 0.011%, respectively, between CH251 and PTB1 (Supplementary Note online). In either case, the conclusion we first reported has not changed: that the diversity of the chimpanzee Y chromosome is much lower than expected, as pioneered by Stone *et al.*<sup>4</sup> and reconfirmed by the independent analysis of Hughes *et al.*<sup>2</sup> using their different strategy.

Yoko Kuroki<sup>1</sup>, Todd D Taylor<sup>1</sup>, Hideki Noguchi<sup>1,2</sup>, Takehiko Ito<sup>3</sup>, Atsushi Toyoda<sup>1</sup>,

Yoshiyuki Sakaki<sup>1</sup> & Asao Fujiyama<sup>1</sup>

<sup>1</sup>RIKEN Genomic Sciences Center, Yokohama, Kanagawa 230-0045, Japan. <sup>2</sup>Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa 277-0882, Japan. <sup>3</sup>Mitsubishi Research Institute, Tokyo 100-8141, Japan.

Note: Supplementary information is available on the Nature Genetics website.

1. Kuroki, Y. *et al.* *Nat. Genet.* **38**, 158–167 (2006).
2. Hughes, J.F. *et al.* *Nature* **437**, 100–103 (2005).
3. The International Chimpanzee Chromosome 22 Consortium. *Nature* **429**, 382–388 (2004).
4. Stone, A.C. *et al.* *Proc. Natl. Acad. Sci. USA* **99**, 43–48 (2002).