

CONCLUSION

After detailed examination of the criticisms presented by Elhaik *et al*, we show that there are both technical and conceptual flaws that undermine their claims. While not central to the arguments of Elhaik *et al*, there are multiple additional problems in their manuscript, some of which we discuss in the Supplementary Note. However, we do wish to point out that the supposed quotation, cited as personal communication FLM, was entirely fabricated, and we have placed the full set of email correspondence between FL Mendez and E Elhaik on our website http://hammerlab.biosci.arizona.edu/supplementary_data.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Fernando L Mendez^{1,2}, Krishna R Veeramah^{1,3}, Mark G Thomas⁴,
Tatiana M Karafet¹ and Michael F Hammer^{*,1}

¹ARL Division of Biotechnology, University of Arizona,
Tucson, AZ, USA;

²Department of Genetics, Stanford University, Stanford, CA, USA;

³Department of Ecology and Evolution, Stony Brook University,
Stony Brook, NY, USA;

⁴Research Department of Genetics, Evolution and Environment,
University College London, London, UK
E-mail: mfh@email.arizona.edu

- 1 Elhaik E, Tatarinova TV, Klyosov AA, Graur D: The 'extremely ancient' chromosome that isn't: a forensic bioinformatic investigation of Albert Perry's X-degenerate portion of the Y chromosome. *Eur J Hum Genet* 2014; **22**: 1111–1116.
- 2 Mendez FL, Krahn T, Schrack B *et al*: An African American paternal lineage adds an extremely ancient root to the human Y chromosome phylogenetic tree. *Am J Hum Genet* 2013; **92**: 454–459.

- 3 Blum MG, Jakobsson M: Deep divergences of human gene trees and models of human origins. *Mol Biol Evol* 2011; **28**: 889–898.
- 4 McDougall I, Brown FH, Fleagle JG: Stratigraphic placement and age of modern humans from Kibish, Ethiopia. *Nature* 2005; **433**: 733–736.
- 5 Thomson R, Pritchard JK, Shen P, Oefner PJ, Feldman MW: Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. *Proc Natl Acad Sci USA* 2000; **97**: 7360–7365.
- 6 Xue Y, Wang Q, Long Q *et al*: Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. *Curr Biol* 2009; **19**: 1453–1457.
- 7 Francalacci P, Morelli L, Angius A *et al*: Low-pass DNA sequencing of 1200 Sardinians reconstructs European Y-chromosome phylogeny. *Science* 2013; **341**: 565–569.
- 8 Cruciani F, Trombetta B, Massaia A, Destro-Bisol G, Sellitto D, Scozzari R: A revised root for the human Y chromosomal phylogenetic tree: the origin of patrilineal diversity in Africa. *Am J Hum Genet* 2011; **88**: 814–818.
- 9 Wei W, Ayub Q, Chen Y *et al*: A calibrated human Y-chromosomal phylogeny based on resequencing. *Genome Res* 2013; **23**: 388–395.
- 10 Poznik GD, Henn BM, Yee MC *et al*: Sequencing Y chromosomes resolves discrepancy in time to common ancestor of males versus females. *Science* 2013; **341**: 562–565.
- 11 Scozzari R, Massaia A, Trombetta B *et al*: An unbiased resource of novel SNP markers provides a new chronology for the human Y chromosome and reveals a deep phylogenetic structure in Africa. *Genome Res* 2014; **24**: 535–544.
- 12 Pink CJ, Swaminathan SK, Dunham I, Rogers J, Ward A, Hurst LD: Evidence that replication-associated mutation alone does not explain between-chromosome differences in substitution rates. *Genome Biol Evol* 2009; **1**: 13–22.
- 13 Taylor J, Tyekucheva S, Zody M, Chiaromonte F, Makova KD: Strong and weak male mutation bias at different sites in the primate genomes: insights from the human-chimpanzee comparison. *Mol Biol Evol* 2006; **23**: 565–573.
- 14 Wilson Sayres MA, Venditti C, Pagel M, Makova KD: Do variations in substitution rates and male mutation bias correlate with life-history traits? A study of 32 mammalian genomes. *Evolution* 2011; **65**: 2800–2815.
- 15 Miyata T, Hayashida H, Kuma K, Mitsuyasu K, Yasunaga T: Male-driven molecular evolution: a model and nucleotide sequence analysis. *Cold Spring Harb Symp Quant Biol* 1987; **52**: 863–867.
- 16 Durrett R: *DNA Sequence Evolution*, 2nd edn. Springer, 2008.
- 17 Birky CW Jr, Walsh JB: Effects of linkage on rates of molecular evolution. *Proc Natl Acad Sci USA* 1988; **85**: 6414–6418.
- 18 Fenner JN: Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. *Am J Phys Anthropol* 2005; **128**: 415–423.
- 19 Casella G, Berger RL: *Statistical Inference*, 2nd edn. Brooks/Cole, 2001.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)

Reply to Mendez *et al*: the 'extremely ancient' chromosome that still isn't

European Journal of Human Genetics (2015) **23**, 567–568;
doi:10.1038/ejhg.2014.227; published online 15 October 2014

Earlier this year, we discovered that an extreme age estimate for a Y chromosomal haplotype (237 000–581 000 years ago) by Mendez *et al*¹ was based on analytical choices that consistently inflated its value.²

As stated in our original criticism,² estimating divergence time is not different, in principle, from estimating the time it takes two cars traveling in opposite directions at known speeds to reach a certain distance from each other. The time inferences will be overestimated if the distance between the two cars is overestimated, or if the speed of either car is

underestimated. Similarly, a divergence time estimate will seem larger than the actual divergence time if the genetic distances between sequences are overestimated and/or the rates of substitution are underestimated.

Let us consider a very simple estimation model for the time of divergence,

$$t = \frac{d}{2r} \quad (1)$$

where t is the divergence time, d is the genetic distance, and r is the substitution rate per unit time. To overestimate t , one needs to overestimate d and/or underestimate r . d is usually estimated by dividing the number of differences between two sequences, n , by the length of the aligned sequences, l , and correcting for multiple hits and the like

$$d = \frac{n}{2l} \quad (2)$$

d can, thus, be overestimated by either overestimating n or underestimating l . The unit time for r is years. However, r is often derived from data on number of substitutions per generation. r can, thus, be overestimated by assuming that the generation time, t_g , is larger than it really is.

In selecting values for d , r , n , l , and t_g , Mendez *et al*¹ consistently and without exception chose values that led to overestimating the time of divergence.

In Elhaik *et al*,² we discussed many such choices. In the following we will focus on two choices left unexplained by Mendez *et al*.³ The first choice concerns the substitution rate used in the calculation of the TMRCA. Using an estimate based on Y-chromosome substitution rate (1×10^{-9} substitutions per nucleotide per year)⁴ we can calculate divergence times of $43/240\,000/10^{-9} \approx 179\,000$ years and $45/180\,000/10^{-9} \approx 250\,000$ years, for an average of 214 500 years, very similar to the TMRCA obtained using a likelihood-based method: 209 500 (95% CI: 168 000–257 400) years.² Not surprisingly, by employing an autosomally derived value of 0.617×10^{-9} as the mutation rate constant, which is 1.6 times smaller, Mendez *et al*¹ obtained a divergence time 1.6 times higher than that estimate of 290 000–404 000 years, with an average value of 347 000 years. More appropriate choices would have resulted in a much lower estimate. Mendez *et al*¹ other choices, such as the unprecedented 40 years for human generation time, resulted in overestimating the time of divergence by 20–130%.

The second choice concerns the irregular and questionable comparison of mutation numbers based on sequences of unequal lengths. Mendez *et al*³ compared 240 000 bases of the A00 Y-chromosome that contained 43 mutations with 180 000 bases of the A0 Y-chromosome that contained 45 mutations. In other words, they used data from two segments, in which one segment was smaller than the other by about 25%. In response to Mendez *et al*'s³ allegations of 'misunderstanding of population genetic theory,' we challenge the authors to come up with one example in the evolutionary literature in which the branches on a phylogenetic tree were estimated by using pairwise distances based on alignments of different lengths. We note that textbooks in

molecular evolution (for example, Graur and Li⁵) specifically caution against such practices.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Thomas Krahn for his comments.

Eran Elhaik^{*,1}, Tatiana V Tatarinova², Anatole A Klyosov³ and Dan Graur⁴

¹Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK;

²Department of Pediatrics, Children's Hospital Los Angeles and Keck School of Medicine, University of Southern California, Los Angeles, CA, USA;

³The Academy of DNA Genealogy, Newton, MA, USA;

⁴Department of Biology & Biochemistry, University of Houston, Houston, TX, USA

E-mail: e.elhaik@sheffield.ac.uk

1 Mendez FL, Krahn T, Schrack B *et al*: An African American paternal lineage adds an extremely ancient root to the human Y chromosome phylogenetic tree. *Am J Hum Genet* 2013; **92**: 454–459.

2 Elhaik E, Tatarinova TV, Klyosov AA, Graur D: The 'extremely ancient' chromosome that isn't: a forensic bioinformatic investigation of Albert Perry's X-degenerate portion of the Y chromosome. *Eur J Hum Genet* 2014; **22**: 1111–1116.

3 Mendez FL, Veeramah KR, Thomas MG, Karafet TM, Hammer MF: Reply to "The 'extremely ancient' chromosome that isn't" by Elhaik *et al*. *Am J Hum Genet* 2014.

4 Xue Y, Wang Q, Long Q *et al*: Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. *Curr Biol* 2009; **19**: 1453–1457.

5 Graur D, Li W-H: *Fundamentals of Molecular Evolution*. Sinauer Associates: Sunderland, MA, USA, 2000.