6 My-old hybridization among human ancestors separated by ~4.5 My of evolution as inferred from analysis of mtDNA pseudogenes (NUMTs).

Konstantin Popadin¹, Konstantin Gunbin², Leon Peshkin³, Sofia Anness³, Zoe Fleischmann⁴, Natasha Markuzon⁵, Yevgenya Kratybskaya⁶, Rebecca Ackermann⁶, and Konstantin Khrapko⁴ (kkrapko@gmail.com)

¹ University of Lausanne, Switzerland, ² Novosibirsk State University, Russia, ³ Harvard Medical School, Boston USA, ⁴ Northeastern University, Boston USA, ⁵ Draper Laboratory, Boston, USA, ⁶ University of Cape Town, South Africa

Abstract:
Increasingly, human evolution is being tied to genetic exchange between divergent lineages (e.g., Neanderthals, Denisovans). However, less is known about genetic exchange during human evolution (pre-1Ma) and between more divergent lineages (a few My apart).

We use the analysis of nuclear pseudogenes of mtDNA ("NUMTs") to look for earlier, more distant hybridization events. NUMTs are fragments of mtDNA that were inserted into our nuclear genome (a few hundred of them have been permanently inserted into our nuclear genome over the last 60 My). NUMTs are considered "mtDNA fossils", as they preserve sequences of ancient mtDNA and thus carry unique information about ancestral populations. Our analysis of a NUMT sequence shared by humans, chimpanzees, and gorillas, implies that, around the time of divergence between humans and chimpanzees, our ancestors interbred with hominins whose mtDNA is that time had diverged as much as ~4.5My from human-to-be mtDNA, strongly implying distant hybridization. Analysis of other NUMTs suggests that such events might have occurred more than once. Reassuringly, interspecies hybridization across such large genetic distances have been observed in other primates, e.g. baboons and colobines. We hypothesize that hybridization might have provided the diversity that promoted evolution of our lineage, and could explain the apparent ascic morphology of some fossils of that epoch.


Fig. 1. Main observation:
The primary subject of this study is a NUMTs on human chromosome 5, which is shared with chimpanzee and gorilla. A joint phylogenetic tree of the three NUMTs variants and the seven mtDNA sequences reveals a very long NUMT stem and short branches of the NUMT subtree, in drastic contrast to the proportions of the mtDNA sub-tree. As expected for a dysfunctional sequence, mutations in the outer pseudogene branches (colored blue) contain a significantly higher proportion of non-synonymous changes (p<0.0005) than mutations of mtDNA branches (green). Intriguingly, mutations of the NUMT stem contain a very high proportion of synonymous changes, as if mutations of mtDNA branches were mixed with mutations of NUMT branches at 75-25 ratio (therefore colored teal).

Fig. 2. Interpretation:
We hypothesize that mixed synonymy of mutations of the NUMT stem suggests that NUMT sequence has been evolving as a functional mitochondrial genome, then turned into a NUMT, which then was inherited by the Homo, Pan, and Gorilla. The number of mutations in the stem is equivalent to ~4.5 My of divergence. Thus, mtDNA lineage leading to the NUMT should have been diverging for 4.5My from the human-to-be mtDNA. This implies that this pseudogene should have been created in a hominin (now extinct) that at that time had been diverged by about 4.5My of evolution from the hominin that at that time carried our mtDNA lineage. In order for this pseudogene and our mtDNA to end up in the same body, these two hominids should have mated. The large divergence implies that such mating was a distant interspecies (or even inter-generic) hybridization.

Fig. 3. Large ancestral N, is not an explanation:
A potential alternative explanation of the high divergence of the mtDNA precursor of the NUMT from human-to-be mtDNA could be that the high effective population size N, of the ancestral population. In this scenario, genetic heterogeneity in population is so large that a divergent individual could be a regular member of the population, rather than an intruder from a distant species. Indeed, the ancestral N, is believed to have been larger than N, of modern apes. To explore how much mtDNA divergence depends on N, we asked (Fig 3) the observed maximal mtDNA divergence in present ape populations vs. N, of the same populations (from Prado-Martinez et al., 2013). Strikingly, there is no correlation between N, and mtDNA divergence. Moreover, even in the largest populations (e.g. common chimpan, P.t.rogl., N,=30,000, which is close to anticipated ancestral N,=50,000), mtDNA divergence is drastically lower than divergence of the NUMT precursor (red diamond), which discounts large N, as an alternative explanation.

Fig. 4. Quantitation:
The estimated divergence of the mitochondrial portion of NUMT stem (representing the divergence of the hypotetical extinct hominin involved in the hybridization) lands in between the divergence between closely related species (e.g. Homo and Pan) and divergences between genera.

Fig. 5. Test for branch length artifacts: short NUMT branches do not distort our estimates
We were concerned that unusually short outer branches of the pseudogene subtree might have artificially forced the ML algorithm to extend the NUMT stem in compensation. We therefore added extra mutations (transitions) into the pseudogene branches which made them essentially undistinguishable from the normal mitochondrial branches by any parameters. Reassuringly, extension of branch lengths essentially did not affect the topology and topography of the remaining tree.

Fig. 6. Speculation: Hybridization generated diversity which sparked hominin speciation
Of note, even a much less distant recent hybridization (~1My) between two baboon species produced a remarkable diversity of cranium sizes and shapes as seen in Fig 6 (Ackerman 2006). It is thus tempting to speculate that the more dramatic hybridization across 4.5 My of divergence around the time of human/chimpanzee separation generated the extra diversity that augmented the separation of the hominin lineage from the apes.

References:
