

# The Y Chromosome and the Origin of All of Us (Men)

Svante Pääbo

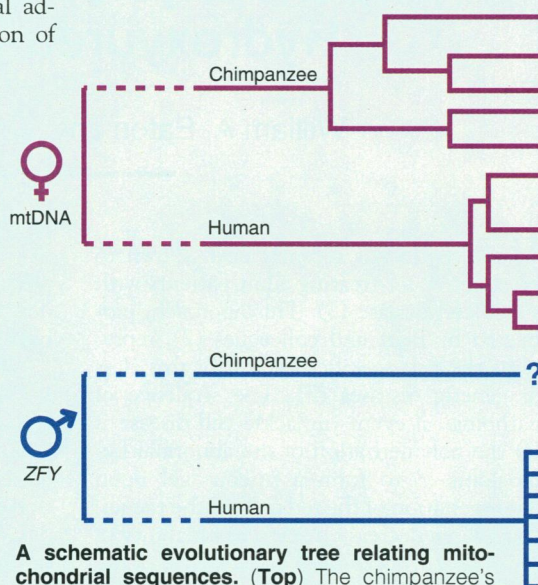
Recently, perhaps even more has been learned about human evolution in molecular genetics laboratories than at archaeological excavations or in museum collections. In this issue of *Science*, Dorit and co-workers (1) report the latest contribution of molecular genetics to unraveling our past.

To study human history, geneticists have traditionally examined the frequencies of genetic variants (or alleles)—usually crudely identified by the electrophoretic mobility of their gene products. Technological advances now allow direct determination of the nucleotide sequences of alleles in a population. We can say with assurance not only that two alleles are different, but how different they are from each other: Two alleles that differ by 20 substitutions are obviously more distantly related than two alleles that differ by a single substitution. As a consequence, the mathematical treatment of how evolutionary lineages trace their ancestry back over generations has increased in sophistication and prominence.

The number of differences among DNA sequences, or nucleotide diversity, in a species or population is directly proportional to a rather arcane quantity that population geneticists refer to as the effective population size ( $N_e$ ). Other things being equal, a smaller population will contain less nucleotide diversity. What makes  $N_e$  particularly interesting is that not only does the current population size matter but also that  $N_e$  is affected by what happened to a population in the past. A reduction and subsequent expansion in population size, such as would occur in a speciation event, will cause a reduction of  $N_e$  and thus of nucleotide diversity. Other factors can also reduce  $N_e$ . For example, if one variant at a genetic locus becomes favored by Darwinian selection, then this variant will sweep through the population, dragging along with it all linked sequences on the chromosome. To distinguish such selection from a past reduction in population size as a cause for reductions in genetic diversity is one of the tasks around which the “genetic archae-

ology” of our species currently revolves.

Some parts of the genome are easier to tap for information than others, and one of the best is mitochondrial DNA. Mitochondrial DNA (mtDNA) is endowed with an evolutionary rate that is higher than that of most regions in the nuclear genome. Thus, with a reasonable amount of sequencing effort, enough variation can be observed to allow the relatedness of mtDNAs to be estimated with some confidence. Furthermore,



**A schematic evolutionary tree relating mitochondrial sequences. (Top)** The chimpanzee's mitochondrial gene pool is about four times deeper than that of the human. **(Bottom)** For the *ZFY* gene, no variation is observed among 38 male humans, and the situation among chimpanzees remains to be studied.

the mtDNA is, for all practical purposes, inherited only from mother to offspring and does not recombine. It therefore gives a picture of our maternal history unadulterated by what the men have been doing. It is not surprising that mtDNA diversity was first applied to the biggest question of all—where do we come from (2, 3)? The answer: from Africa and quite recently, some 100,000 to 200,000 years ago.

The first conclusion—Africa—was based on two findings. First, Africans have more mitochondrial sequence diversity among them than other populations, and the diversity between Africans and others is much smaller than that found within Africa. Thus, in a genetic sense, everyone on this planet looks like an African. The easiest explanation for this is that humans

originated in Africa, and that what we see genetically in the mtDNA is the effect of smaller groups that have left Africa and colonized various parts of the globe. Another possible explanation, eagerly pointed out by some, is that African populations have historically been larger and thus, because of a larger  $N_e$ , would contain more diversity than others. The second line of evidence that pointed to Africa was phylogenetic trees reconstructed from mtDNA sequences collected from around the globe. When the root of such trees was sought by connecting them to chimpanzee sequences, the root seemed to fall among African sequences. However, reanalyses of the data (4) indicated that there is very little information in the sequences to justify this conclusion, mainly because the chimpanzee sequence is so distantly related to that of humans that most of the information on its relation to human sequences has been erased by substitutions occurring multiple times at the same positions.

However, the recent date of our mitochondrial ancestor is in a sense the really controversial conclusion from these studies. Everyone agrees that we trace our ancestry to *Homo erectus*, who emerged in Africa and from there colonized most of Eurasia about a million years ago or even earlier. What the mitochondrial data seem to show, however, is that we have a much more recent ancestor, one who lived some 100,000 to 200,000 years ago. Those earlier ancestors must then have been replaced by modern humans in a second emigration out of Africa. Thus, evidence from the mtDNA suggests that we are all quite closely related, much more so than, for example, chimpanzees (see figure). However, there are other possible explanations. For example, if a mitochondrial substitution that conferred some substantial advantage in reproductive success or survival occurred some 100,000 years ago, then this substitution could have swept around the Old World and could be presenting us with a picture today that is true for mtDNA but not for the genome in general. Thus, everyone has eagerly awaited dates for the diversity of other regions of the genome that could support or contradict the mitochondrial evidence. On page 1183 of this issue, Dorit and co-workers give us one such piece of evidence (1).

All eyes have been on the non-recombining part of the Y chromosome as the next genomic region to study. This DNA is inherited paternally and thus gives the male side of our history, neatly complementing the mitochondrial picture. Dorit *et al.* have sequenced an intron of the *ZFY* gene on the nonrecombining part of the Y chromosomes of 38 men from around the world. To their surprise, they found no nucleotide difference among them. In

The author is in the Institut of Zoology, University of Munich, P.O. Box 202136, D-80021 Munich, Germany.



contrast, the same sequence shows many differences when compared among humans, chimpanzees, gorillas, and orangutans. Confronted with the mathematically rather awkward situation of dealing with the absence of variation, they calculate the minimum amount of nucleotide diversity that must exist in this intron in order to allow their observation of no variation to lie within the 95% confidence interval of the expected. They then go on to estimate a time to the common ancestor of the Y chromosome by calculations that, surprisingly, do not take the  $N_e$  of males into account. The date they come up with is 270,000 years ago, more recent than can be envisioned for the origin of *Homo erectus*. It thus fits with the recent mitochondrial date of the modern human gene pool.

Could there be other explanations for this observation? Peter Goodfellow and his group (5), as well as others, have studied the evolution of another gene on the Y chromosome, the SRY, which is directly responsible for inducing maleness. They find that when this gene is compared among species, it evolves very rapidly. Some parts of the protein even show such a high number of amino acid replacement substitutions relative to silent substitutions that one must conclude either that part of the protein has no function at all or that its evolution is driven by directional selection—a very rare phenomenon among protein-coding genes. Yet, within species very little variation is found. How could this be? A possible explanation is the absence of recombination on the Y chromosome. Normally, a mutation that is selected will, on its way to fixation, be broken loose from accompanying regions on the chromosome where it initially occurred by recombination. Not so on the Y chromosome. Here, a selected allele will drag with it all the rest of the non-recombining part of the Y chromosome in potentially gigantic selective sweeps. This could homogenize the gene pool within the species, yet allow rapid evolution between species. So does the finding of Dorit *et al.* reflect a lack of recombination on human Y chromosomes rather than say something about our past?

A recent paper by Nachman and Aquadro (6) is of relevance to this question. They have compared the amount of within-species polymorphism among mice for a noncoding region flanking the Sry gene and related it to the amount of between-species divergence for the same sequence. They compared this ratio to the same ratio for the rapidly evolving control region of the mtDNA in the same mice. If any of the two loci were the victims of a selective sweep in the recent past, this would show up in the form of a reduced intraspecific polymorphism. The results show that they cannot

reject the hypothesis that the Sry, and thus the entire nonrecombining part of the Y chromosome, evolves neutrally. Therefore, there seems to be no reason to believe that selection acts on the Y chromosome.

Of course, the mtDNA data, as well as the ZFY data, could both be explained by selective sweeps. However, when two loci rather than one tell the same story, the selectional explanation becomes less likely. But more work is still needed on how to calculate the coalescence time when no variation is observed. It is also of importance to see to what extent the ZFY intron varies among nonhuman primates. Other regions in the human genome, so small that recombination over the time span relevant for human history will not be a major factor,

should also be examined to show whether our recent common ancestry is a genome-wide phenomenon. So far, we have just seen the small beginning of what someone has called the biggest archaeological excavation of all times: the quest into the genome to reveal our past.

## References

1. R. L. Dorit, H. Akashi, W. Gilbert, *Science* **268**, 1183 (1995).
2. R. L. Cann, M. Stoneking, A. C. Wilson, *Nature* **325**, 31 (1987).
3. L. Vigilant, M. Stoneking, H. Harpending, K. Hawkes, A. C. Wilson, *Science* **253**, 1503 (1991).
4. S. B. Hedges, S. Kumar, K. Tamura, M. Stoneking, *ibid.* **255**, 737 (1992).
5. L. S. Whitfield, R. Lovell-Badge, P. N. Goodfellow, *Nature* **364**, 713 (1993).
6. M. W. Nachman and C. F. Aquadro, *Mol. Biol. Evol.* **11**, 539 (1995).

# The Biophysics of Sick Cell Hydroxyurea Therapy

William A. Eaton and James Hofrichter

Hydroxyurea has recently been shown to be effective for treating adult patients with sickle cell disease (1). This approach, pioneered by Platt and colleagues (2), represents the first specific therapy for this class of genetic diseases (3). The sequence of pathological events in sickle cell disease is (i) the polymerization of the abnormal hemoglobin S to form a viscous gel upon deoxygenation of the red cells in the tissues, (ii) a large decrease in the deformability of the cells, (iii) occlusion of the vessels of the microcirculation by the stiffened cells, (iv) insufficient oxygen delivery, and finally, (v) tissue damage, which may cause episodes of severe pain known as a "sickle cell crisis." This sequence is less likely in patients with an unusually high fraction of "F cells." F cells contain about 20% fetal hemoglobin (hemoglobin F) and 80% hemoglobin S, whereas "S cells" contain only hemoglobin S. The principal therapeutic effect of hydroxyurea in sickle cell patients is believed to be the increase in the fraction of F cells, from about 30% to about 50% (1). Although the biochemical mechanism by which hydroxyurea stimulates hemoglobin F synthesis is not yet known, the biophysical mechanism of how hemoglobin F has such a profound inhibitory effect on polymerization in F cells is well understood. The

story of this mechanism is an example of physics and physical chemistry providing direct and deep insights into the molecular pathology of a human disease (4).

The net effect of hemoglobin F in F cells is simply to dilute the hemoglobin S. How does such a small decrease in hemoglobin S concentration have a therapeutic effect? Kinetics hold the key (5). Upon deoxygenation of sickle red cells, there is a pronounced delay before the explosive appearance of polymer (see figure). If the delay time is less than the ~1 s transit time of a red cell through the microcirculation, polymerization will occur while the cell is squeezing through small vessels, and blockage may occur (see figure). If the delay time is longer than the 10 to 20 s that it takes the cell to return to the lungs to be reoxygenated, polymerization will not occur. Although polymerization takes place in almost all sickle cells at venous oxygen pressures at equilibrium in vitro, there is no polymerization in the vast majority of cells in vivo because the delay time is longer than 10 to 20 s (6).

The delay time is incredibly sensitive to solution conditions, inversely proportional to about the 30th power of the initial hemoglobin S concentration (5). In contrast, ordinary chemical reaction rates depend on the first or second power of the reactant concentration. Consequently, a small decrease in the concentration of hemoglobin S has a large effect on the delay time. Decreasing the intracellular hemoglobin S concentration to delay polymerization is

The authors are in the Laboratory of Chemical Physics, National Institute of Diabetes, Digestive and Kidney Diseases, Building 5, National Institutes of Health, Bethesda, MD 20892-0520, USA.