The history of the Y chromosome in man

Jennifer F. Hughes & David C. Page

Studies of the Y chromosome over the past few decades have opened a window into the history of our species, through the reconstruction and exploitation of a patrilineal (Y-genealogical) tree based on several hundred single nucleotide variants (SNVs). A new study validates, refines, and extends this tree by incorporating >65,000 Y-linked variants identified in 1,244 men representing worldwide diversity.

Despite its puny size, the Y is among the most storied of human chromosomes. The refined genealogical tree of modern human Y chromosomes reported by Chris Tyler-Smith, Carlos Bustamante and colleagues in this issue[1] extends this narrative legacy, and ensures that it will continue well into the future.

Why the Y?

The Y chromosome is home to the testis-determining gene *Sry*, which causes fetuses to develop as anatomic males. The Y and its meiotic partner, the X, thereby qualify as sex chromosomes. But here the ironies begin. Across 95% of its length (all but its pseudoautosomal tips) the Y chromosome abstains from sexual recombination – the exchange of genes – with the X. This has been true throughout the history (and long before the origin) of our species. Thus the great bulk of the Y chromosome has been transmitted clonally – asexually – from father to son down the generations, its content shaped only by mutation and selection, without swapping genetic material with a partner during meiotic cell division, as occurs on all other nuclear chromosomes,

including the X. The isolationist behavior of the Y chromosome unfortunately resulted in its decline, both in size and in gene repertoire, but it also turned the Y into a powerful device for recording the migratory and demographic history of the males of our species. Random, stable mutations, such as single nucleotide variants (SNVs), accumulate over time across the Y chromosome, and particular collections of these mutations are permanently linked together because of the Y chromosome's asexual transmission. A set of human males who share a particular collection of Y-chromosome mutations is called a haplogroup, and those shared mutations can be traced back to a common patrilineal ancestor. Y-chromosome transmission across generations is analogous to an (asexually) expanding yeast colony, where random mutations accumulate in some cells and are passed on to progeny during cell division. Some such mutations in yeast have striking phenotypic consequences, and their origin and spread can be traced through the colony (Fig. 1). In the same way, Y-chromosome SNV patterns gathered from men across the world have been used to delineate a detailed phylogenetic tree showing relationships among extant Y chromosomes, which gives us insight into our species' history[2-4].

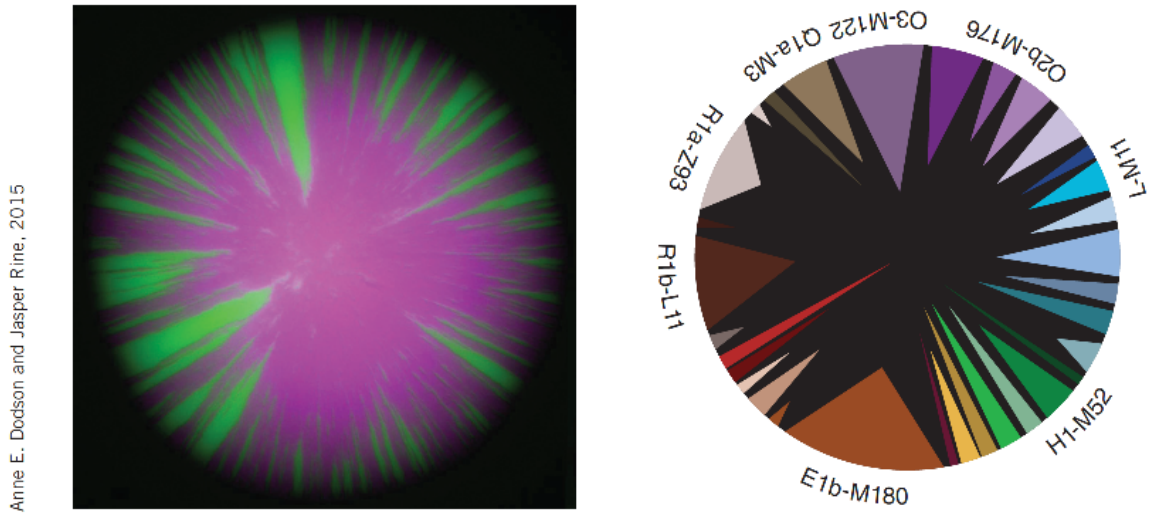

Figure 1. Clonal expansion of mutants in a yeast colony (at left) and among a geographically diverse collection of human Y chromosomes (at right; an artist's reinterpretation of Fig. 2 from Poznik et al.). Pink-to-green mutants in the yeast propagate as clonal sectors (credit: Anne E. Dodson and Jasper Rine, http://elifesciences.org/content/4/e05007v2).

Tracing human history

The new study by Poznik et al. utilizes the unprecedented dataset of whole genome sequences generated by the 1000 Genomes Project[5], which includes individuals selected from 26 geographically diverse human populations. In total, the authors analyzed Y-chromosome sequences from 1,244 men. They restricted their analysis to single-copy regions of the Y chromosome, which equate to ~10 million base pairs or roughly half of the Y chromosome's total euchromatic sequence. This strategy ensured that sequence variants they identified represented polymorphisms (differences between orthologous sequences in individuals) rather than differences between paralogous copies of long (>10 thousand base pairs), nearly identical repeats, which are prominent on the Y chromosome. Because the Y is present in only one copy per cell, the authors had to contend with relatively low sequence coverage (~4.3 fold). The authors were therefore conservative in their approach, and their account of >65,000 variants (including SNVs, insertions, deletions, and short tandem repeat length variation) is likely an underestimate of the true level of variation among the Y chromosomes studied.

Based on the distribution of SNVs, the authors constructed a phylogenetic tree showing the paternal lineages connecting all of the 1,244 men studied. The structure of this new tree mirrors previous trees, which were based on 100-fold fewer SNVs[3,4,6]. However, the wealth of information provided by both the large number of Y chromosomes and variants included in the current study further refine the tree and provide new insight into human population dynamics. For example, the expanded analysis reveals a new megagroup, which originated ~55,000 years ago and encompasses nearly all non-African males, and an ancient clade within haplogroup H, which is prominent in South Asia. At least eight major population expansions are evident in the new phylogenetic tree, and the timing of a number of these expansions correlates with notable events in human history.

Unlike earlier studies that focused on a set of previously ascertained SNVs, the new study is unbiased because it catalogs all variation detected across a large expanse of the Y

chromosome. This approach enables a robust method for calculating dates of bifurcations, or population splits, within the tree. The authors estimate the date for the most recent common ancestor of all Y chromosomes, which is rooted in Africa, to be ~190,000 years, and for all non-African Y chromosomes to be ~76,000 years. These dates are considerably older than those reported in previous, but less expansive, Y-chromosome sequencing efforts[7-9]; the discrepancies are due in part to differences in the Y-chromosome mutation rate used for calibration.

Importantly, this new dataset provides abundant information for conducting future studies of Y-chromosome genealogy. However, tracking the inheritance of the Y chromosome provides a masculinized, and in some instances misleading, view of the history of our species. For example, reproductive success can differ dramatically among individuals due to societal or political factors, which could lead to the predominance of one particular Y chromosome in a population. This observation of reduced Y-chromosome diversity could be misinterpreted as evidence of a severe population bottleneck. The most famous example of this phenomenon is seen in the legacy of Ghenghis Khan, whose Y-chromosome descendants comprise ~8% of men in a large region of Asia[10], and another study found additional examples of the spread of such "super Ys" in recent human history[11]. The Y chromosome provides just one piece of a complex puzzle of human history, with mitochondrial DNA, which is maternally inherited, and autosomal loci filling in the rest. Just as this present study was not possible a decade ago, 10 years from now, we will likely have effectively unlimited access to genome-wide variation from across the globe, allowing a more encompassing reconstruction of our species' history.

References

1. Poznik, G. D. *Nat Genet* **48**, 593-599 (2016).
2. Underhill, P. A. et al. *Nat Genet* **26**, 358-361 (2000).
3. The Y Chromosome Consortium. *Genome Res* **12**, 339-348 (2002).
4. Karafet, T. M. et al. *Genome Res* **18**, 830-838 (2008).
5. Consortium, T. G. P. *Nature* **526**, 68-74 (2015).
6. Jobling, M. et al. *Nature Reviews: Genetics* **4**, 598-612 (2003).
7. Hallast, P. et al. *Mol Biol Evol* **32**, 661-673 (2015).
8. Poznik, G. D. et al. *Science* **341**, 562-565 (2013).

9. Wei, W. et al. *Genome Res* **23**, 388-395 (2013).
10. Zerjal, T. et al. *Am J Hum Genet* **72**, 717-721 (2003).
11. Balaresque, P. et al. *Eur J Hum Genet* **23**, 1413-1422 (2015).

Author information

Jennifer F. Hughes is at the Whitehead Institute, Cambridge, Massachusetts, USA.

David C. Page is at the Whitehead Institute, Howard Hughes Medical Institute, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

Competing financial interests

The authors declare no competing financial interests.

Corresponding author

Correspondence to: Jennifer F. Hughes (jhughes@wi.mit.edu)