Published in final edited form as:

Nat Genet. 2009 January ; 41(1): 8–10. doi:10.1038/ng0109-8.

Evaluating signatures of sex-specific processes in the human genome

Carlos D. Bustamante and

Department of Biological Statistics and Computational Biology, Cornell University, 102 Weill Hall, Ithaca, New York 14850, USA

Sohini Ramachandran

Society of Fellows and Department of Organismic and Evolutionary Biology, Harvard University, 4100 Biological Laboratories, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA. e-mail: cdb28@cornell.edu

Abstract

Comparing levels of genetic variation between the X chromosome and autosomes can reveal the different demographic histories of males and females of a species. Taking this approach, two new studies report that the effective population sizes of men and women differ, but they disagree as to which sex outnumbered the other.

Evolutionary forces and social practices that differentially impact males and females can lead to divergent patterns of genetic variation among the autosomes, X chromosome, Y chromosome and the mitochondrial genome^{1,2}. For example, populations in which a few males father a disproportionate number of offspring show low levels of Y-chromosomal genetic diversity. Likewise, societies in which females migrate to their husbands' homes after marriage (patrilocal) experience more gene flow of X chromosomes and mtDNA^{3,4}. Investigating genomic signatures of sex-biased processes has relied on comparing the nonrecombining Y chromosomes and mitochondrial genomes⁵, and it is hoped that the availability of genome-level data will provide a much finer-scale view of both patterns and processes^{6–8}. On page 66 of the current issue, David Reich and colleagues⁹ report evidence for accelerated genetic drift on the X chromosome (that is, smaller effective population size, N_e , of females) during human migrations out of Africa, based on finding disproportionately low sequence variation and high SNP frequency differentiation of the X chromosome relative to that of the autosomes in non-African populations. Their results are in contrast to a recent paper by Michael Hammer and colleagues, who with a different study design estimated an unexpectedly high X-chromosomal diversity¹⁰.

More men on the move?

Keinan *et al.*⁹ compared genetic diversity among 130,000 X-chromosomal and autosomal SNPs genotyped in individuals of Northern European, West African and East Asian ancestry^{11,12}. They use a statistic Q that compares rates of genetic drift of X-linked and autosomal SNPs and has an expected value of 0.75 when the effective population size of males and females is equal. Comparing the Northern European and East Asian samples, they find that Q is statistically indistinguishable from the expectation of 0.75 (ref. ⁹). However, comparing either non-African population with the West African samples, Q drops significantly below 0.75, which suggests faster genetic drift for the X chromosome in populations outside of Africa. Keinan *et al.*⁹ also compared large amounts of sequence data found in public databases for five individuals of Northern European ancestry, four of East Asian ancestry and five of West African ancestry. They find ratios of pairwise sequence diversity close to 0.75 in

the West African population and below 0.75 outside of Africa (Fig. 1), which supports their SNP analysis. They demonstrate with simulations that demographic models with a smaller effective number of mating females in non-African populations are necessary to explain their results. Some sex-biased forces that could produce this pattern include recurrent long-range male migration from African sources or a shorter generation time for females.

But the story may not be so clear. In a study with a similar goal of comparing X-chromosomal and autosomal genetic variation, Michael Hammer and colleagues¹⁰ analyzed over 210 kilobases of DNA sequence data from across 20 independent regions on the X chromosome and 20 independent regions on the autosomes, sequenced in each of 90 individuals from six geographically diverse populations (Fig. 1a)¹⁰. They estimate the ratio of effective population size on the X chromosome (N_X) and autosomes (N_A) from two summary statistics of the data—the observed divergence between human and orangutan (used to account for variation in mutation rate among chromosomes) and the number of segregating sites in the sample (that is, SNPs in the data). They observe N_X/N_A is above 0.75 for all six surveyed populations, and in three cases it is significantly higher than 0.75 with no systematic out-of-Africa effect. They therefore conclude that the effective population size of females is larger than that of males. They also use simulations to demonstrate that sex-biased processes are needed to explain their observation, and they suggest that larger variance in male reproductive success relative to females (such as would occur under polygyny) may explain the data.

Finding common ground

In order to reconcile these findings, we compared a common summary statistic, π (pairwise nucleotide diversity), of the data reported by the two groups (Fig. 1). While the estimated autosomal nucleotide diversity (π_{aut}) is very similar for both studies ($\approx 0.12\%$ per site or 1 SNP per 900 base pairs in African populations, and $\approx 0.08\%$ or 1 SNP per 1,250 base pairs in European and East Asian populations), the estimated X-chromosomal nucleotide diversity (π_X) differs substantially (Fig. 1a). Keinan *et al.*⁹ report approximately 0.72, 0.46 and 0.41 SNPs per kilobase of X chromosome in West African, North European and East Asian populations, respectively, whereas Hammer *et al.*¹⁰ find about 40–50% more nucleotide variation in each of their samples for the closest comparable populations (0.99, 0.71 and 0.58 for Mandenka, Han Chinese and Basque, respectively). Both studies effectively use π_X/π_{aut} to estimate N_e for males and females, so the higher overall levels of X-chromosomal diversity in Hammer *et al.*¹⁰ explain part of the discrepancy. It appears that the rest of the discrepancy is explained by different normalizations for background mutation rate differences between the X chromosome and autosomes (Hammer *et al.*¹⁰ used human-orangutan divergence and Keinan *et al.*⁹ used human-macaque divergence) (Fig. 1b).

While we are able to identify that the papers differ in their estimates of X-chromosomal genetic diversity and X-to-autosome mutational bias, how and why these differences arise remain open questions. Part of the answer may lie in the advantages and limitations of each study design. Hammer *et al.*¹⁰ studied population-level processes with direct resequencing among a relatively large number of individuals (which we advocate), but they investigated a limited number of genomic regions, selected to be far from genes and with high recombination rates to minimize possible confounding effects of natural selection. Keinan *et al.*⁹ used two different sources of genome-wide data, starting with SNP data, which is inherently more difficult to interpret due to unknown ascertainment biases in the data, and we commend their attempt to control for many such biases. They demonstrate further support for the observed signature of faster genetic drift on the X chromosome by comparing nucleotide diversity among pairs of chromosomes from each population. This approach, however, is less powerful than comparing the full spectrum SNP frequencies one observes in deep resequence data. Encouragingly, the genetic diversity estimates reported by Keinan *et al.*⁹ are nearly identical to those recently

reported by Bentley *et al.*¹³ for a pair of fully sequenced X chromosomes of a Northern European female (0.47%).

In order to address continuing questions on the nature of sex-biased processes, full genome sequencing of large numbers of individuals sampled from diverse populations will be needed. The upcoming 1,000 Genomes Project (<http://www.1000genomes.org/>), for example, will provide orders of magnitude more data for these types of analyses. We share the enthusiasm of the population genetics community that this will bring the potential for resolving continuing questions regarding how human history and cultural practices have shaped global patterns of genomic diversity.

References

1. Wilkins JF. *Curr Opin Genet Dev* 2006;16:611–617. [PubMed: 17067791]
2. Schaffner SF. *Nat Rev Genet* 2004;5:43–51. [PubMed: 14708015]
3. Seielstad MT, Minch E, Cavalli-Sforza LL. *Nat Genet* 1998;20:278–280. [PubMed: 9806547]
4. Wilder JA, Kingan SB, Mobasher Z, Pilkington MM, Hammer MF. *Nat Genet* 2004;36:1122–1125. [PubMed: 15378061]
5. Ségurel L, et al. *PLoS Genet* 2008;4:e1000200. [PubMed: 18818760]
6. Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, Stoneking M. *Nat Genet* 2001;29:20–21. [PubMed: 11528385]
7. Wilkins JF, Marlowe FW. *Bioessays* 2006;28:290–300. [PubMed: 16479583]
8. Ramachandran S, Rosenberg NA, Feldman MW, Wakeley J. *Theor Popul Biol* 2008;74:291–301. [PubMed: 18817799]
9. Keinan A, Mullikin JC, Patterson N, Reich D. *Nat Genet* 2009;41:66–70. [PubMed: 19098910]
10. Hammer MF, Mendez FL, Cox MP, Woerner AE, Wall JD. *PLoS Genet* 2008;4:e1000202. [PubMed: 18818765]
11. International HapMap Consortium. *Nature* 2005;437:1299–1320. [PubMed: 16255080]
12. International HapMap Consortium. *Nature* 2007;449:851–861. [PubMed: 17943122]
13. Bentley DR, et al. *Nature* 2008;456:53–59. [PubMed: 18987734]

Figure 1. Contrasting patterns of X-chromosomal and autosomal diversity. Data are from Keinan *et al.*⁹ and Hammer *et al.*¹⁰. **(a)** Pairwise sequence diversity values for the X chromosome (π_X) and autosomes (π_{aut}). **(b)** Ratios of X-chromosomal to autosomal sequence diversities (π_X/π_{aut}), with and without divergence corrections used by each study. A horizontal line is drawn at 0.75, the expected value for the ratio. YRI, Yoruban (Ibadan, Nigeria); CHB, Han Chinese (Beijing, China); JPT, Japanese (Tokyo, Japan); CEU, CEPH (Utah residents with ancestry from Northern and Western Europe)^{11,12}.