# Sex-Averaged Recombination and Mutation Rates on the X Chromosome: A Comment on Labuda et al.

*To the Editor:* A recent paper by Labuda et al.[1] used patterns of linkage disequilibrium (LD) among SNPs from the HapMap data on the X chromosome and the autosomes to estimate the female-to-male breeding ratio in human populations ($\beta = N_f/N_m$). This approach was of considerable interest to us because two recent papers[2,3] using SNP diversity and frequency patterns to study sex-biased demography differed in their conclusion as to whether the effective population size of the X chromosome was larger than expected. A larger than expected effective population size on the X chromosome could be due to a larger female than male effective population size ($\beta > 1$). Because neither of the previous studies used information contained within LD patterns, the study of Labuda et al.,[1] in principle, could provide independent estimates of $\beta$. They find evidence that $\beta$ is slightly larger than 1 but still smaller than the value reported by Hammer et al.[2] Thus, Labuda et al.[1] concluded that there is little evidence for polygyny, or a larger female than male effective population size, throughout human history. However, errors in their analytical derivations affect most of their analyses, and correction of these errors leads to different conclusions.

In deriving Equation 4, Labuda et al.[1] state that the sex-averaged recombination rate on the X chromosome, $r_X$, depends on the female-to-male breeding ratio of the population through the expression $r_X = (2\beta/(1+2\beta))r_{fX}$, in which $r_{fX}$ is the female recombination rate. However, $r_X = (2/3)r_{fX}$ and is independent of $\beta$ because each offspring is produced from a male-female mating, regardless of the sex ratio in the population. Therefore, because recombination on the X chromosome can occur only in females ($r_{mX} = 0$), only two of the three potentially transmitted X chromosomes can be the product of a recombination event. Deviations from an equal number of breeding males and females in the population will change the relationship between the effective population sizes of the X chromosome ($N_{eX}$) and the autosomes ($N_{eA}$), but will not change the fact that each mating will still consist of a single male parent and a single female parent (Figure 1), keeping $r_X = (2/3)r_{fX}$. Thus, the authors' expression essentially double-corrects for unequal male-female population sizes. The correct expression ($r_X = (2/3)r_{fX}$) has been previously derived (reviewed in [4]) and has also been used to interpret differences in patterns of genetic variation on the X chromosome and autosomes in *Drosophila*.[5,6] The expression for the sex-averaged recombination rate on the X chromosome is the same for humans and *Drosophila* because, in both species, it does not recombine in males.

Using the correct equation for $r_X$, Equation 4 of Labuda et al.[1] should read

$$\rho_X = \frac{8}{3} N_{eX} r_{fx}\,.$$

Then, it follows that the X chromosome-to-autosome ratio of population recombination rates (Equation 7) should be

$$\frac{\rho_X}{\rho_A} = \frac{r_{fX}}{r_A}\frac{2N_{eX}}{3N_{eA}} = \frac{r_{fX}}{r_A}\frac{3(\beta+1)}{4(\beta+2)},$$

in which $r_A$ is the sex-averaged recombination rate on the autosomes. The ratio of the normalized X chromosome recombination rate to the normalized autosomal recombination rate ($R$) defined in Equation 8 then becomes

$$R = \frac{\rho_X}{\rho_A}\frac{r_A}{r_{fX}} = \frac{2N_{eX}}{3N_{eA}} = \frac{3(\beta+1)}{4(\beta+2)}.$$

The breeding ratio as a function of $R$ (captured in Equation 9) is

$$\beta = \frac{8R-3}{3-4R}.$$

Figure 2 shows the population recombination rate ratio (solid blue curve) along with the ratio computed from Equation 8 of Labuda et al.[1] (dotted blue curve). Equation 8 of Labuda et al.[1] underpredicts $R$ when $\beta$ is low (an excess of breeding males) and overpredicts $R$ when $\beta$ is high (an excess of breeding females).

Given that it appears that the error in the derivations of Labuda et al.[1] has a substantial impact on $R$ (Figure 2), we reanalyzed the data presented in Table 1 of Labuda et al.[1] from the three HapMap populations. We calculated $\beta$ from the estimates of $R$ from Labuda et al.,[1] using the corrected version of Equation 9. The corrected Equation 9 results in larger estimates of $\beta$ than those reported in Table 1 of the original paper[1] (see Table 1 in this paper). For example, in YRI, $\beta = 2.63$, as compared to 1.42 before correction. In terms of $N_{eX}/N_{eA}$, the corrected equation gives a ratio of 0.882 in YRI instead of 0.796 reported by Labuda et al.[1] These larger estimates of $\beta$ and $N_{eX}/N_{eA}$ from the HapMap CEU and YRI populations are consistent with the estimates reported in Hammer et al.[2] and support the claim of an excess of breeding females in human history. Incidentally, although we follow Labuda et al.[1] in reporting results in terms of $\beta$, we note that $N_{eX}/N_{eA}$ is a more robust statistic and that deriving $\beta$ from $N_{eX}/N_{eA}$ introduces the restrictive assumptions of discrete non-overlapping generations and a Poisson distribution of offspring.[7–9]

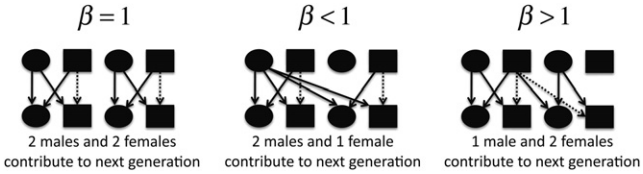


**Figure 1. Illustration of the Biological Model Underlying Different Breeding Ratios in the Population**
A single generation of reproduction is shown, in which an equal number of males and females reproduce ($\beta = 1$), more males than females reproduce ($\beta < 1$), and more females than males reproduce ($\beta > 1$). Solid arrows denote the transmission of a copy of the autosomal genome in addition to an X chromosome. Dotted arrows denote the transmission of only an autosomal genome. Importantly, an X chromosome spends 2/3 of its time in females, regardless of $\beta$, as evidenced by four out of six copies of it being inherited from a female in each of the three panels.

Labuda et al.[1] then used their estimates of $\beta$ from the LD patterns in the HapMap data combined with diversity and divergence levels on the X chromosome and autosomes to estimate the ratio of male germline mutations to female germline mutations ($\alpha$). The expression that the authors derived for the sex-averaged X chromosome mutation rate ($\mu_X$) depends on $\beta$. For the same reasons described above with regard to $r_X$, $\mu_X$ is independent of $\beta$ as well. Corrected expressions for Equations A2–A6 of Labuda et al. are presented in Appendix S1, available online. Importantly, when the correct expressions are used, the ratio of X chromosome-to-autosome diversity ($\Theta_X/\Theta_A$) follows a monotonically increasing function of $\beta$ for all values of $\alpha$ (Figure S1), rather than the complex pattern shown in Figure 2 of Labuda et al.[1] The corrected expressions, corrected estimates of $\beta$ (Table 1), and the estimates of $\Theta_X/\Theta_A$ from Table S2 of Labuda et al.[1] provide estimates of $\alpha$ between 4.95 and 22.43. These estimates are higher than those obtained by Labuda et al.,[1] though estimates

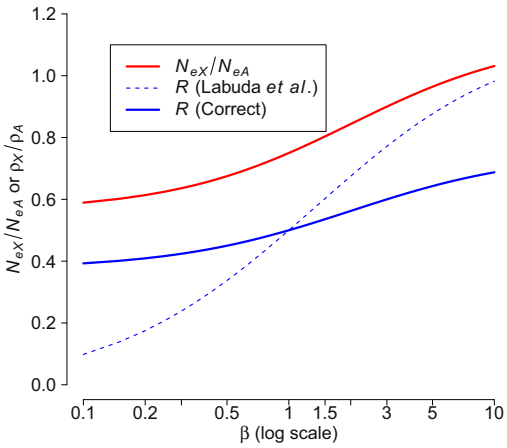


**Figure 2. Ratio of Effective Population Sizes, $N_{eX}/N_{eA}$, and Population Recombination Rates, $R$, as a Function of the Breeding Ratio in the Population, β**
The dotted blue curve denotes $R$ calculated from Equation 8 of Labuda et al.[1] The solid blue curve denotes $R$ calculated from our corrected equation (see text).

**Table 1. Original and Corrected Estimates of $\beta$ and $N_{eX}/N_{eA}$ from Table 1 of Labuda et al.**

| | $\rho_X$[a] | $\rho_A$[a] | $R$[a] | Original β[a] | Corrected β | Original $N_{eX}/N_{eA}$[a] | Corrected $N_{eX}/N_{eA}$ |
|---|---|---|---|---|---|---|---|
| YRI | 0.264 | 0.449 | 0.588 | 1.42 | 2.63 | 0.796 | 0.882 |
| CEU | 0.136 | 0.237 | 0.574 | 1.34 | 2.27 | 0.788 | 0.861 |
| CHB, JPT | 0.158 | 0.301 | 0.525 | 1.10 | 1.33 | 0.763 | 0.787 |

[a] Reproduced from Table 1 of Labuda et al.[1]

of $\alpha$ equal to 5 have been previously noted in humans.[4,10,11] The highest estimate of $\alpha$ is from the YRI population, which has the largest estimate of $\beta$. The reliability of this estimate is unclear, because $\Theta_X/\Theta_A$ may differ across populations[3] and the data used by Labuda et al.[1] do not account for this. Furthermore, it is not clear that estimates of $\beta$ from LD-based summary statistics can be used to obtain reliable estimates of mutational parameters, given that Labuda et al.'s work[1] and previous work[12] have shown that complex demography can affect SNP diversity and frequency patterns differently than it affects LD patterns.

Labuda et al.[1] also suggested that estimates of $\alpha$ from X chromosome and autosome divergence depend on the sex ratios of the populations involved. However, this is at odds with previous work showing that when ignoring ancestral polymorphism, $\alpha$ can be estimated solely from the X chromosome versus autosome divergence without regard to $\beta$.[4,13,14]

In conclusion, we applaud Labuda et al.'s[1] use of LD-based summary statistics to distinguish between competing complex demographic models. However, errors in their analytical derivations undermine their conclusion that there is little evidence for larger female than male effective population sizes throughout human history. Instead, when the corrected equations presented here are used, their results from some populations are consistent with a female effective population size roughly twice that of males.

Kirk E. Lohmueller,[1,*] Jeremiah D. Degenhardt,[2] and Alon Keinan[2]
[1]Department of Integrative Biology, University of California, Berkeley, CA 94720, USA; [2]Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY 14853, USA
*Correspondence: k.lohmueller@berkeley.edu

## Supplemental Data

Supplemental Data include a Supplemental Appendix and one figure and can be found with this article online at http://www.cell.com/AJHG.

## Acknowledgments

We thank Andrew Clark and Rasmus Nielsen for helpful discussions and comments on previous versions of the manuscript. We

also thank other members of the Nielsen laboratory for helpful discussions and two anonymous reviewers for helpful comments on the manuscript. K.E.L. was supported by a Ruth Kirschstein National Research Service Award from the National Human Genome Research Institute (F32HG005308).

### References

1. Labuda, D., Lefebvre, J.F., Nadeau, P., and Roy-Gagnon, M.H. (2010). Female-to-male breeding ratio in modern humans-an analysis based on historical recombinations. Am. J. Hum. Genet. *86*, 353–363.
2. Hammer, M.F., Mendez, F.L., Cox, M.P., Woerner, A.E., and Wall, J.D. (2008). Sex-biased evolutionary forces shape genomic patterns of human diversity. PLoS Genet. *4*, e1000202.
3. Keinan, A., Mullikin, J.C., Patterson, N., and Reich, D. (2009). Accelerated genetic drift on chromosome X during the human dispersal out of Africa. Nat. Genet. *41*, 66–70.
4. Hedrick, P.W. (2007). Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. Evolution *61*, 2750–2771.
5. Andolfatto, P., and Przeworski, M. (2000). A genome-wide departure from the standard neutral model in natural populations of Drosophila. Genetics *156*, 257–268.
6. Przeworski, M., Wall, J.D., and Andolfatto, P. (2001). Recombination and the frequency spectrum in *Drosophila melanogaster* and *Drosophila simulans*. Mol. Biol. Evol. *18*, 291–298.
7. Caballero, A. (1995). On the effective size of populations with separate sexes, with particular reference to sex-linked genes. Genetics *139*, 1007–1011.
8. Charlesworth, B. (2001). The effect of life-history and mode of inheritance on neutral genetic variability. Genet. Res. *77*, 153–166.
9. Laporte, V., and Charlesworth, B. (2002). Effective population size and population subdivision in demographically structured populations. Genetics *162*, 501–519.
10. Ellegren, H. (2007). Characteristics, causes and evolutionary consequences of male-biased mutation. Proc. Biol. Sci. *274*, 1–10.
11. Makova, K.D., and Li, W.H. (2002). Strong male-driven evolution of DNA sequences in humans and apes. Nature *416*, 624–626.
12. Lohmueller, K.E., Bustamante, C.D., and Clark, A.G. (2009). Methods for human demographic inference using haplotype patterns from genomewide single-nucleotide polymorphism data. Genetics *182*, 217–231.
13. Graur, D., and Li, W.H. (2000). Fundamentals of Molecular Evolution (Sunderland, MA: Sinauer).
14. Miyata, T., Hayashida, H., Kuma, K., Mitsuyasu, K., and Yasunaga, T. (1987). Male-driven molecular evolution: a model and nucleotide sequence analysis. Cold Spring Harb. Symp. Quant. Biol. *52*, 863–867.

DOI 10.1016/j.ajhg.2010.03.021. 

# Response to Lohmueller et al.

*To the Editor:* In this issue, Lohmueller et al. rightly noted that we doubly corrected for unequal male-female population sizes, a mistake that inadvertently perpetuated itself in subsequently derived equations. We are grateful to these authors for pointing out our mistake so quickly and thus helping us to rapidly correct our calculations. We complete the corrections made by Lohmueller et al. in their comment in our Supplemental Data, available online, where we show correct versions of the derived equations and updated resulting figures and tables.

Our mistake led us to underestimate the breeding ratio β. The corrected estimates are greater but still within a range of ratios of the male-to-female reproductive variance encountered in societies characterized as monogamous or serially monogamous, although they also overlap with those characterizing polygyny.[1] Our updated estimates are at the low end of the estimates obtained by Hammer et al., which ranged from 1.8 to 14,[2] and thus do not strongly support the results and conclusions discussed by these authors.

Importantly, in addition to capturing sex differences in the reproductive variance, β can be affected by sex differences in the generation time, by sex-biased migration or inbreeding, as well as by matrilocality or patrilocality and possibly by sex-asymmetric admixture.[3,4] Furthermore, following a population bottleneck, β estimates can be skewed as a result of a faster equilibration of a genetic system of lower effective population size, such as that of the X chromosomes versus the autosomes. Therefore, estimates of β from population-diversity data have to be interpreted in the context of demographic, anthropological, evolutionary, and paleontological evidence.[1,3]

Our estimates of β were derived from the ratio of $N_{eX}/N_{eA}$ estimated from the ratio of the population recombination rates of these chromosomal systems. Lohmueller et al. remarked that $N_{eX}/N_{eA}$ is a more robust statistic than β itself. In addition, focusing first on $N_{eX}/N_{eA}$, it may be easier to partition the distinct contributions of the factors enumerated above to the overall numeric outcome of this ratio in order to eventually extract only the part influenced by the breeding ratio and use it directly to estimate β. This is, however, conditional on the data and the genetic information that can be used to evaluate distinct contributing parameters. Combining information that can be obtained from historical recombinations[3] with that obtained from mutations[2,4,5] should help this task, both in testing population models and in refining the resulting estimates.

Using our new approach, one can extract additional information from the genetic-variability data to confront different estimates obtained independently from the analysis of the mutational diversity and to examine their consistency. Divergence of such estimates prompts additional investigations. For example, the estimate of about 5 of the ratio, α, of the male-to-female mutation rate,