Evolution, 58(7), 2004, pp. 1613–1616

ESTIMATING THE STRENGTH OF SEXUAL SELECTION FROM Y-CHROMOSOME AND MITOCHONDRIAL DNA DIVERSITY

MICHAEL J. WADE[1] AND STEPHEN M. SHUSTER[2]
[1]*Department of Biology, Indiana University, 1001 East Third Street, Bloomington, Indiana 47405*
E-mail: mjwade@indiana.edu
[2]*Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona 86011-5640*
E-mail: stephen.shuster@nau.edu

Abstract.—We show that a sex difference in the opportunity for selection results in sex differences in the strength of random genetic drift and thus creates different patterns of genetic diversity for maternally and paternally inherited haploid genes. We derive the effective population size N_e for a male-limited or female-limited haploid gene in terms of I, the ''opportunity for selection'' or the variance in relative fitness. Because the variance in relative fitness of males can be an order of magnitude larger than that of females, the N_e is much smaller for males than it is for females. We derive both nonequilibrium and equilibrium expressions for F_{ST} in terms of I and show how the portion of I owing to sexual selection, I_{mates}, that is, the variation among males in mate numbers, is a simple function of the F's for cytoplasmic (female inherited) and Y-linked (male inherited) genes. Because multiple, transgenerational data are lacking to apply the nonequilibrium expression, we apply only the equilibrium model to published data on Y chromosome and mitochondrial sequence divergence in *Homo sapiens* to quantify the opportunity for sexual selection. The estimate suggests that sexual selection in humans represents a minimum of 54.8% of total selection, supporting Darwin's proposal that sexual selection has played a significant role in human evolution and the recent proposal regarding a shift from polygamy to monogamy in humans.

Key words.—Effective population size, F_{ST}, haplotype diversity, *Homo sapiens,* opportunity for selection, sex difference in selection, sexual selection.

Received July 30, 2003. Accepted March 31, 2004.

Sexual selection ''. . . depends, not on a struggle for existence, but on a struggle between males for possession of the females; the result is not death of the unsuccessful competitor, but few or no offspring'' (Darwin 1859, p. 88). Sexual selection is one of the strongest known evolutionary forces, responsible, within species, for differences between males and females in coloration, weaponry, and sperm morphology and, among species, for large differences in male phenotypes (Endler 1986; Andersson 1994; Shuster and Wade 2003). As a consequence of reproductive competition, typically some males have many mates whereas many other males have no mates at all. This variation in mate numbers is the cause of sexual selection and in most species makes selection on males several times stronger than selection on females (Wade 1979; Arnold and Wade 1984a,b; Fairbairn and Wilby 2000; see also below). We show how a sex difference in the opportunity for selection results in sex differences in the strength of random genetic drift and creates different patterns of genetic diversity for maternally inherited and paternally inherited haploid genes. We use studies of Y chromosome and mitochondrial sequence divergence in *Homo sapiens* to quantify the opportunity for sexual selection. We estimate that sexual selection in humans represents a minimum of 54.8% of total selection, supporting Darwin's proposal that sexual selection has played a significant role in human evolution.

Differences in the diversity of human Y-linked and mitochondrial markers have been interpreted as evidence for a higher female than male migration rate (Seielstad et al. 1998; Perez-Lezaun et al. 1999) or a more recent expansion of the Y chromosome out of Africa than previously thought (Thomson et al. 2003) while the role of sexual selection has been discounted. Some studies have found more modest reductions in levels of Y-linked diversity (e.g. Nachman 1998). However, these studies used an indirect measure of sexual selection based on the sex ratio of breeding males to females and not on a formal sex-specific measure of the opportunity for selection (e.g., Seielstad et al. 1998) and its effect on genetic drift. The sex-ratio approach used by other authors seriously underestimates sexual selection's effect on Y-chromosome diversity as we show below. Low diversity of Y-chromosome sequence variation is common (Dupanloup et al. 2003). For example, from the observation of a single Y-chromosome haplotype originating in Mongolia approximately 1000 years ago, it has been inferred that the Mongol warlord, Genghis Khan, and his male relatives were the progenitors of nearly 8% of the men now living in the geographic region formerly controlled by the Mongol Empire, which is approximately 1.0% of the world's human population (Zerjal 2003). If true, such accounts indicate that sexual selection in human populations has the potential to be very strong. However, the actual intensity of sexual selection remains unclear.

It has been shown elsewhere (Wade 1979, 1995; Arnold and Wade 1984a,b; Shuster and Wade 2003) that the variance in reproductive fitness among males, V_{male}, is related to the variance in offspring number of females, V_{female}, through the following equation:

$$V_{male} = RV_{female} + (W_{female})^2 V_{mates}, \quad (1)$$

where R is the sex ratio, N_{female}/N_{male}; W_{female} is the mean number of offspring per female; and V_{mates} is variance in mate numbers among males. Although the term V_{mates} measures the variance in male fitness due to differences in mate numbers among males, it is evaluated in terms of the number of offspring males obtain with each mating. Thus, when paternity can be accurately assigned, this method provides a powerful means for quantifying male fitness variance, particularly

when multiple mating by females or cryptic female mate choice leads to differential fertility among males (Shuster and Wade 2003). Because each individual in a sexual population has one father and one mother, the mean fitness of males, W_{male}, must equal RW_{female}, the mean fitness of females multiplied by R, the mean number of mates per male. The term $(W_{female})^2V_{mates}$, in equation (1) can make the variance in male reproductive success as much as 2.0×10^3 times larger than that of females (Shuster and Wade 2003). The greater the variance in offspring numbers, the stronger is the force of random genetic drift (Crow and Kimura 1970, p. 100). Consequently, a sex difference in fitness variance will cause a sex difference in the strength of random genetic drift, which will manifest itself as different levels of diversity in those genes with sex-limited inheritance (as opposed to sex-limited gene expression). Diversity will be lower in the sex with the greater fitness variance because random drift will be stronger in that sex.

The variance in relative fitness from all causes, genetic and environmental, affects random drift. The variance in relative fitness, I, is called the ''opportunity for selection'' (Arnold and Wade 1984a,b; Crow 1989; Shuster and Wade 2003), and it equals the variance in absolute fitness divided by the mean fitness squared. Because the opportunity for selection measures the increase in average fitness of the breeding parents *relative* to that of the parent population before selection, it is a measure of the opportunity for selection. Dividing equation (1) by mean male fitness squared, $RW_{female}{}^2$ (Wade 1979; Arnold and Wade 1984a,b; Shuster and Wade 2003), gives:

$$I_{male} = (1/R)I_{female} + I_{mates}, \tag{2}$$

where the ratio $V_{female}/W_{female}{}^2$ is I_{female}, the opportunity for selection on females, and I_{mates} is V_{mates}/R^2, the variance in mate numbers among males divided by R^2, the square of mean number of mates per male. I_{mates} is the opportunity for sexual selection among males; that is, the variance in relative male fitness that results from the variance among males in mate numbers (Wade 1979; Arnold and Wade 1984a,b; Shuster and Wade 2003). It includes mating as well as nonmating males. When R is unity, I_{mates} equals $(I_{male} - I_{female})$, the sex difference in the opportunity for selection in many circumstances (Wade 1979, 1995; Arnold and Wade 1984a,b; Shuster and Wade 2003). Moreover, the value of I_{male} can be much greater than that of I_{female}. For example, in the northern water snake, *Nerodia sipedon*, paternity analysis using microsatellite markers reveals that I_{male} is 70 times greater than I_{female} (Gibbs and Weatherhead 2001). Similarly, in the marine isopod, *Paracerceis sculpta*, I_{male} is more than 20 times greater than I_{female} (Shuster and Wade 2003). The magnitude of the sex difference in sex-limited gene diversity will depend on the magnitude of the sex difference in the opportunity for selection, with predictable patterns in species with conventional as well as with reversed sexual roles. Thus, this method is also useful for quantifying the opportunity for sexual selection in sex-role reversed species. For example, in the broad-nosed pipefish, *Nerophis ophidion*, males carry a single brood of offspring, but females may mate with up to three males. In this species, I_{female} is three times greater than I_{male} (Berglund et al. 1989).

Selection, by definition, affects the variance in offspring numbers. It reduces N_e, the effective population size (Seielstad et al. 1998; Perez-Lezaun et al. 1999), and thereby enhances genetic drift. The greater the variance in relative fitness, the smaller N_e becomes and the stronger is random genetic drift, all else being equal (i.e., selection, migration, and mutation). Consequently, sex differences in the operation of drift cause differences in the amount of variation in maternally and paternally inherited genes. For a neutral haploid gene with uniparental inheritance, where W is mean fitness (Crow 1958; Crow and Kimura 1970; Wade 1996):

$$N_e = (N - [1/W])/(1 + [V/W^2] - [1/W]). \tag{3}$$

When the average number of gametes per parent is one (i.e., replacement; Wright 1938), this reduces to $(N - 1)/(I)$, recognizing that I is the ratio (V/W^2). When W is very large, the expression reduces to $N/(I + 1)$. Thus, for $1 \leq W < \infty$, we have $(N/[I + 1]) < N_e < ([N - 1]/I)$. In the following derivation, we use the upper bound on N_e. (Use of the lower bound also leads to equation [7], below, when k [defined below] is 1.0 and equation [7] remains an underestimate for $k > 1$.) Throughout our derivation, we assume as in most previous studies that the mitochondrial and Y-linked variants are neutral and that direct selection on these haplotypes is negligible. There are instances in which this assumption fails and, in the interpretation of data in the context of these equations, this assumption should be investigated (e.g., Nachman 1998) to support the interpretation based on the neutrality assumption.

Because I_{male} is often an order of magnitude larger than I_{female} (Gibbs and Weatherhead 2001; Shuster and Wade 2003; see above), the effective population size of a male-inherited gene, like those on the Y chromosome, can be very much smaller than that of a female-inherited gene, like those of the mitochondria. The harmonic mean sex ratio of breeding adults (Seielstad et al. 1998) is often used to evaluate the reduction in the N_e of males caused by polygyny relative to the N_e of biparentally inherited nuclear genes. This approach is appropriate only under the assumption of a random distribution of offspring numbers, such that the ratio, (V/W), is one for both sexes. Note that when the variance in fitness equals the mean $(W = V)$, as expected in the absence of selection, I is one and N_e equals $(N - 1)$, or approximately N. The existence of a large sex difference in the variance in relative fitness invalidates this assumption and the use of this approach greatly underestimates the effect of sexual selection on the difference in N_e between males and females.

The effective population size, N_e, determines the genetic diversity maintained at equilibrium within a population by the forces of random genetic drift, mutation (rate μ), and migration (rate m). For a haploid gene with the standard assumption that $m \gg \mu$, the general equation (Wright 1951) for the equilibrium probability of identity by descent, or the among-population variation, is

$$F \approx 1/(2N_em + 1). \tag{4}$$

Substituting equation (3) for N_e and rearranging, we obtain

$$F \approx I/(2[N - 1][m] + I). \tag{5}$$

TABLE 1. Estimates of F_{ST} (or the related parameters, Φ_{ST} and R_{ST}) for nucleotide sequences located on the Y chromosome and on mitochondrial DNA (mtDNA) in human populations. The ratio of sexual selection to total selection is estimated using equation (7) and assuming that k equals 1. (If $k > 1$, then the tabled values are underestimates).

Region	Parameter	Y chromosome	mtDNA	% sexual selection	Data source
(a) Europe	Φ_{ST}	0.187	0.057	58.4	Kittles et al. 1999
	R_{ST}	0.090	0.043	37.5	Kittles et al. 1999
(b) Native American	Φ_{ST}	0.137	0.151	−5.7	Kittles et al. 1999
	R_{ST}	0.190	0.080	45.9	Kittles et al. 1999
	Φ_{ST}	0.349	0.230		Kittles et al. 1999
(c1) Combined a and b	R_{ST}	0.178	0.054	28.4	Kittles et al. 1999
(c2)	F_{ST}	0.185	0.062	58.3	Jorde et al. 2000
(d) Europe, Asia, Africa	F_{ST}	0.645	0.186	54.9	Seielstad et al. 1999
(e) Europe				77.7	
Average of c1, c2, d, and e			54.8		

For maternally inherited mitochondrial haplotypes, F_{mt}, at equilibrium is

$$F_{mt} \approx I_{female}/(2[N_{female} - 1][m_{female}] + I_{female}) \quad (6a)$$

and the corresponding equation for Y-linked, paternally inherited haplotypes is

$$F_Y \approx I_{male}/(2[N_{male} - 1][m_{male}] + I_{male}). \quad (6b)$$

When R is unity ($N_{female}/N_{male} = 1$) and the sex-specific migration rates are $m_{male} = km_{female}$, (where k is a constant of proportionality), then the ratio of the difference, $\{(1/F_{mt}) - (1/F_Y)\}$, to the sum, $\{(1/F_Y) + (1/F_{mt}) - 2\}$, is given by

$$\{(1/F_{mt}) - (1/F_Y)\}/\{(1/F_y) + (1/F_{mt}) - 2\} = (I_{mates} - [k - 1]I_{female})/(I_{male} + kI_{female}). \quad (7)$$

When $k = 1$, equation (7) estimates the proportion of total selection represented by sexual selection, namely, $I_{mates}/(I_{male} + I_{female})$. When $k > 1$, the migration rate of females exceeds that of males, as has been reported for humans (Seielstad et al. 1998; Perez-Lezaun et al. 1999). In this case, equation (7) is an *underestimate* of the proportion of sexual selection. Thus, the minimum strength of sexual selection (I_{mates}) relative to total selection ($I_{male} + I_{female}$) can be estimated from equation (7) using Y-linked and mitochondrial haplotype diversities, even when k exceeds 1.

Nonequilibrium Gene Diversity

For humans, given high migration rates and temporally and spatially variable selection, it might be better to avoid the assumption of genetic equilibrium (cf. Whitlock and McCauley 1999). In this nonequilibrium case, we can use the standard recursion equation (ignoring migration for the single generation between t and $[t + 1]$)

$$F(t + 1) = (1/N_e) + (1 - [1/N_e])F(t), \quad (8a)$$

$$\Delta F = F(t + 1) - F(t) = (1/N_e) - (F[t]/N_e). \quad (8b)$$

This equation describes both Y-linked and mitochondrial diversity as long as N_e is understood to be the number of effective males or females, respectively. If we assume that N_e is large and F is small, then terms of magnitude (F/N_e) will be very small compared to $(1/N_e)$. Ignoring such terms, and substituting the upper bound, $(N - 1)/I$ for N_e, we find

$$\Delta F \sim (1/N_e) = I/(N - 1). \quad (9)$$

If we assume that N_{male} equals N_{female} as we did above, then the ratio

$$\frac{(\Delta F_{male} - \Delta F_{female})}{(\Delta F_{male} + \Delta F_{female})} \sim \frac{(I_{male} - I_{female})}{(I_{male} + I_{female})} \quad (10a)$$

$$= \frac{I_{mates}}{(I_{male} + I_{female})}. \quad (10b)$$

To apply this nonequilibrium formulation, we require data on mitochondrial and Y-linked diversity from the same population(s) for two consecutive generations and such data are not presently available from the published literature. For that reason, we must turn to the equilibrium solution, with all the caveats that requires.

The best available data for applying our formulas come from our own species, *Homo sapiens*, in which the distribution of genetic diversity has been studied for mitochondrial, Y chromosomes, and nuclear autosomal genes in the same populations (Wright 1951; Seielstad et al. 1998; Perez-Lezaun et al. 1999; Santos et al. 1999; Dupanloup et al. 2003; Hammer et al. 2003). Estimating sexual selection using the nonequilibrium formula (eq. 10b) requires estimates of F_y and F_{mt} from two sequential generations excluding migrants. No such data currently exist for our species or any other. Thus, we use the reported values (Wright 1951; Poloni et al. 1997; Seielstad et al. 1998; Kittles et al. 1999; Perez-Lezaun et al. 1999; Santos et al. 1999; Jorde et al. 2000) of F_{ST} (or related parameters, Φ_{ST} and R_{ST}) given in Table 1 and apply the equilibrium equation (eq. 7). Using these data and equation (7), we obtain combined estimates indicating that sexual selection represents an average of 54.8% (range 28.4–77.7%) of all selection in *H. sapiens*.

Humans are sexually dimorphic in several traits. Compared to females, human males, on average, are taller and heavier than females, with more facial and body hair, deeper voices, larger brains, higher juvenile mortality, and shorter longevity (Barkow et al. 1992). Human size dimorphism is consistent with an evolutionary history of more intense reproductive competition between males than females, and sexual selection has been accorded a large and controversial role in human evolution with respect to warfare (Alexander 1971), human intelligence (Alexander 1971; Ridley 1993), bipedalism (Dar-

win 1874; Parker 1987), language (Darwin 1874; Ridley 1993) and culture (Barkow et al. 1992). Greater male fitness variance than female fitness variance has been documented in modern human populations (Mulder 1988). However, until now, in humans, or for that matter, in any other animal population, no quantitative method has been available for estimating the *historical* intensity of sexual selection relative to total selection that is captured in the patterns of diversity of sex-limited genes. Following his catalog of human traits possibly influenced by sexual selection, Darwin (1874; p. 600) lamented, ''The views here advanced, on the part which sexual selection has played in the history of man, want scientific precision.'' Our finding that more than half of all selection in humans has been sexual selection supports Darwin's original inference (1874) of its central role in human evolution. More importantly, our method provides a novel means for estimating the relative opportunity for sexual selection in any species for which sex-specific molecular genetic markers exist.

Acknowledgments

We thank C. Lively, J. Demuth, J. Moorad, E. Omana, J. Hey, and anonymous reviewers for comments on the manuscript, and National Science Foundation grants DEB-0212582 and DBI-0243914 for support.

Literature Cited

Alexander, R. D. 1971. The search for an evolutionary philosophy of man. Proc. R. Soc. Victoria: Melbourne 84:99–120.

Andersson, M. 1994. Sexual selection. Princeton Univ. Press, Princeton, NJ.

Arnold, S. J., and M. J. Wade. 1984a. On the measurement of selection in natural and laboratory populations: theory. Evolution 38:709–719.

———. 1984b. On the measurement of selection in natural and laboratory populations: application. Evolution 38:720–733.

Barkow, J. H., L. Cosmides, and J. Tooby, eds. 1992. The adapted mind: evolutionary psychology and the generation of culture. Oxford Univ. Press, Oxford, U.K.

Berglund, A., G. Rosenquist, and I. Svensson. 1989. Reproductive success of females limited by males in two pipefish species. Am. Nat. 133:506–516.

Crow, J. F. 1958. Some possibilities for measuring selection intensities in man. Hum. Biol. 30:1–13.

———. 1989. Fitness variation in natural populations. Pp. 91–97 *in* W. G. Hill and T. F. C. McKay, eds. Evolution and animal breeding. C.A.B. International, Wallingford, U.K.

Crow, J. F., and M. Kimura. 1970. An introduction to population genetic theory. Harper and Row, New York.

Darwin, C. R. 1859. On the origin of species. Facsimile of the first edition, 1964, Harvard Univ. Press, Cambridge, MA.

———. 1874. The descent of man and selection in relation to sex. Rand, McNally and Co., New York. Republished by Gale Research Company, 1974, Book Tower, Detroit, MI.

Dupanloup, I., L. Pereira, G. Bertorelle, F. Calafell, M. J. Prata, A. Amorim, and G. Barbujani. 2003. A recent shift from polygyny to monogamy in humans is suggested by the analysis of worldwide Y-chromosome diversity. J. Mol. Evol. 57:85–97.

Endler, J. 1986. Natural selection in the wild. Princeton Univ. Press, Princeton, NJ.

Fairbairn, D. J., and A. E. Wilby. 2000. Inequality of opportunity: measuring the potential for sexual selection. Evol. Ecol. Res. 3: 667–686.

Gibbs, H. L., and P. J. Weatherhead. 2001. Insights into population ecology and sexual selection in snakes through the application of DNA-based genetic markers. J. Hered. 92:173–179.

Hammer, M. F., F. Blackmer, D. Garrigan, M. W. Nachman, and J. A. Wilder. 2003. Human population structure and its effects on sampling Y chromosome sequence variation. Genetics 164: 1495–1509.

Jorde, L. B., W. S. Watkins, M. J. Bamshad, M. E. Dixon, C. E. Ricker, M. T. Seielstad, and M. A. Batzer. 2000. The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. Am. J. Hum. Genet. 66: 979–988.

Kittles, R. A., A. W. Bergen, M. Urbanek, M. Virkkunen, M. Linnoila, D. Goldman, and J. C. Long. 1999. Autosomal, mitochondrial, and Y chromosome DNA variation in Finland: evidence for a male-specific bottleneck. Am. J. Phys. Anthropote 108:381–399.

Mulder, M. B. 1988. Reproductive success in three Kipsigis cohorts. Pp. 419–435. *in* T. E. Clutton-Brock, ed. Reproductive success: studies of individual variation in contrasting breeding systems. Univ. of Chicago Press, Chicago, IL.

Nachman, M. W. 1998. Y chromosome variation of mice and men. Mol. Biol. Evol. 15:1744–1750.

Parker, S. T. 1987. A sexual selection model for hominid evolution. Hum. Evol. 2:235–253.

Perez-Lezaun, A., F. Calafell, D. Comas, E. Mateu, E. Bosch, R. Martinez-Arias, J. Clarimon, G. Fiori, D. Luiselli, F. Facchini, D. Petterner, and J. Bertranpetit. 1999. Sex-specific migration patterns in Central Asian populations, revealed by analysis of Y-chromosome short tandem repeats and mtDNA. Am. J. Hum. Genet. 65:208–219.

Poloni, E. S., O. Semino, G. Passarino, A. S. Santachiara-Benerecetti, I. Dupanloup, A. Langaney, and L. Excoffier. 1997. Human genetic affinities for Ychromosome P49a,f/TaqI haplotypes show strong correspondence with linguistics. Am. J. Hum. Genet. 61:1015–1035.

Ridley, M. 1993. The red queen: sex and the evolution of human nature. Macmillan, New York.

Santos, F. R., A. Pandya, C. Tyler-Smith, S. D. J. Pena, M. Schanfield, W. R. Leonard, L. Osipova, M. H. Crawford, and R. J. Mitchell. 1999. The central Siberian origin for Native American Y chromosomes. Am. J. Hum. Genet. 64:619–628.

Seielstad, M. T., E. Minch, and L. L. Cavalli-Sforza. 1998. Genetic evidence for a higher female migration rate in humans. Nat. Genet. 20:278–280.

Seielstad, M. T., E. Bekele, M. Ibrahim, A. Toure, and M. Traore. 1999. A view of modern human origins from Y chromosome microsatellite variation. Genome Res. 9:558–567.

Shuster, S. M., and M. J. Wade. 2003. Mating systems and strategies. Princeton Univ. Press, Princeton, NJ.

Thomson, R., J. K. Pritchard, P. Shen, P. J. Oefner, and M. W. Feldman. 2003. Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. Proc. Natl. Acad. Sci. USA 13:7360–7365.

Wade, M. J. 1979. Sexual selection and variance in reproductive success. Am. Nat. 114:742–747.

———. 1995. Mean crowding and sexual selection in resource polygynous mating systems. Evol. Ecol. 9:118–124.

——— 1996. Adaptation in subdivided populations: kin selection and interdemic selection. Pp. 381–405 *in* R. Rose and G. Lauder, eds. Evolutionary biology and adaptation. Sinauer Associates, Sunderland, MA.

Whitlock, M., and D. E. McCauley. 1999. Indirect measures of gene flow and migration: F-st does not equal 1/(4Nm+1). Heredity 82:117–125.

Wright, S. 1938. Size of population and breeding structure in relation to evolution. Science 87:430–431.

———. 1951. The genetical structure of populations. Ann. Eugenics 15:323–354.

Zerjal, T. Y. Xue, G. Bertorelle, R. S. Wells, W. Bao, S. Zhu, R. Qamar, Q. Ayub, A. Mohyuddin, S. Fu, P. Li, and many others. 2003. The genetic legacy of the Mongols. Am. J. Hum. Genet. 72:717–712.

Corresponding Editor: J. Hey