Beyond 2/3 and 1/3: The Complex Signatures of Sex-Biased Admixture on the X Chromosome

Amy Goldberg[1] and Noah A. Rosenberg
Department of Biology, Stanford University, Stanford, California 94305-5020

COMPARISONS of the X chromosome and the autosomes provide a strategy for understanding the history of sex-biased demography (Ramachandran *et al.* 2004, 2008; Wilkins and Marlowe 2006; Hammer *et al.* 2008, 2010; Bustamante and Ramachandran 2009; Keinan *et al.* 2009; Casto *et al.* 2010; Emery *et al.* 2010; Labuda *et al.* 2010; Lambert *et al.* 2010; Gottipati *et al.* 2011; Heyer *et al.* 2012; Arbiza *et al.* 2014). Unlike the autosomes, the X chromosome follows a sex-specific inheritance pattern, with females inheriting two copies, one from the mother and one from the father, and males inheriting only a single copy from the mother. As a consequence, demographic differences between females and males—in such phenomena as the breeding population size, the variance of reproductive success, and migration rates—can be studied by examining differences in patterns of genetic variation between X chromosomes and autosomes.

Many of the best-known cases of sex-biased patterns in human demography involve recently admixed populations (Parra *et al.* 1998, 2001; Bedoya *et al.* 2006; Wilkins 2006; Lind *et al.* 2007; Berniell-Lee *et al.* 2008; Wang *et al.* 2008; Stefflova *et al.* 2009; Bryc *et al.* 2010a,b, 2015; Quintana-Murci *et al.* 2010; Beleza *et al.* 2012; Moreno-Estrada *et al.* 2013; Verdu *et al.* 2013; Marcheco-Teruel *et al.* 2014; Salzano and Sans 2014). During the formation of such populations, sex-biased admixture occurs if one or more of the source populations contributes different fractions of the females and males to the admixed group. What patterns are expected for X chromosomes and autosomes in an admixed population formed through a sex-biased admixture process? An initial hypothesis, reflecting the fact that in a population with equally many females and males, two-thirds of the X chromosomes are in females and one-third are in males, is that if s_1^f is the fraction of females originating from population 1 and s_1^m is the fraction of males originating from population 1, then the

fraction of ancestry from population 1 for a site on the X chromosome is

$$H_1^X = \frac{2}{3}s_1^f + \frac{1}{3}s_1^m \qquad (1)$$

(Lind *et al.* 2007). This simple linear combination has been used to estimate the sex-specific contributions from females and males to African-American and Latino populations (Lind *et al.* 2007; Bryc *et al.* 2015). As we will show, however, it presumes a very specific history for the admixture process, a history often not viewed as reasonable for practical admixture scenarios.

Here, we extend a two-sex mechanistic admixture model (Goldberg *et al.* 2014) to incorporate the genetic signatures of sex-specific admixture patterns on the X chromosome. We derive a recursive expression for the expectation of the X-chromosomal admixture fraction as a function of sex-specific admixture parameters, demonstrating that the X-chromosomal admixture is obtained from a more complex formula than in the simple 2/3 and 1/3 weighting (Equation 1). The limiting mean X-chromosomal admixture is a predictable function of female and male contributions from the source populations, but among cases we consider, has the 2/3 and 1/3 weighting only for a single admixture event that takes place at a single point in time. During the approach to this limit, the behavior of the mean X-chromosomal admixture is dependent on the time since admixture. For a single admixture event and for constant ongoing admixture, we characterize the difference between the limit and the mean X-chromosomal admixture under the model as a function of time. We reinterpret admixture patterns in recently admixed African-Americans, demonstrating that consideration of both the admixture model and the time since admixture is important for estimating sex-specific admixture contributions.

A Mechanistic, Sex-Specific Model for Admixture Histories

We follow the discrete-time model and notation of Verdu and Rosenberg (2011) and Goldberg *et al.* (2014), in which two source populations, S_1 and S_2, contribute to an admixed population, H (Figure 1). For each population, female and male contributions are considered separately. At generation g, the contribution of sex δ, with $\delta \in \{f, m\}$, from source S_α, with $\alpha \in \{1, 2\}$, is $s_{\alpha,g}^\delta$. Corresponding contributions from H are denoted h_g^δ. Thus, for example, s_g^δ denotes the fraction in generation g of individuals of sex δ originating in the previous generation in population S_α. The female and male contributions from S_1, H, and S_2 at generation g are $s_{1,g}^f$, $s_{1,g}^m$, h_g^f, h_g^m, $s_{2,g}^f$, and $s_{2,g}^m$.

We recall key relations among the sex-specific parameters (Goldberg *et al.* 2014, equations 1–6). First, the total contribution from each population is the mean of female and male contributions: $s_{\alpha,g} = (s_{\alpha,g}^f + s_{\alpha,g}^m)/2$ and $h_g = (h_g^f + h_g^m)/2$. Also, as each parameter is a probability, the total female and male contributions separately sum to one: $s_{1,g}^f + h_g^f + s_{2,g}^f = 1$
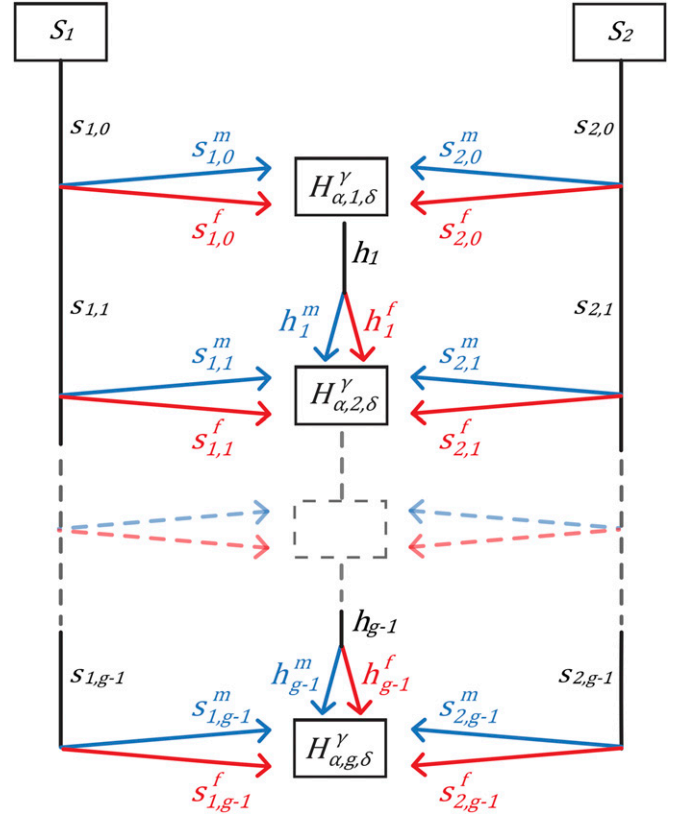
Figure 1 Sex-specific mechanistic model of admixture. At each generation g, females and males from each of two source populations, S_1 and S_2, contribute to the admixed population, H. Contributions can vary in time. The fraction of admixture from source population $\alpha \in \{1, 2\}$ for chromosomal type $\gamma \in \{A, X\}$ in an admixed individual of sex $\delta \in \{f, m\}$ is $H_{\alpha,g,\delta}^\gamma$. This model, by considering different chromosomal types, generalizes the model of Goldberg *et al.* (2014).

and $s_{1,g}^m + h_g^m + s_{2,g}^m = 1$. For $g = 0$, the admixed population does not yet exist, so that $h_0 = h_0^f = h_0^m = 0$ and $s_{1,0}^f + s_{2,0}^f = s_{1,0}^m + s_{2,0}^m = 1$.

Under this general framework, we define sex bias as a difference in the contributions from one or more source populations to the admixed population. That is, an admixture process is considered sex-biased if $s_{1,g}^f \neq s_{1,g}^m$ or $s_{2,g}^f \neq s_{2,g}^m$ or both. Thus, sex bias involves differences between females and males entering from a specified source population, rather than a comparison of the sex-specific parameters between the two source populations; after the founding, it is possible that sex bias can occur in one source but not in the other.

Considering the three populations (S_1, H, S_2) from which the parents of an individual from the admixed population can originate, we have nine possible ordered parental pairings (Table 1). As in Goldberg *et al.* (2014), we study the random variable representing the admixture fraction of a random individual in the admixed population. Whereas Goldberg *et al.* (2014) examined admixture only for the autosomes, here we also study the X chromosome. We define $H_{1,g,\delta}^\gamma$ as the admixture fraction of chromosomal type γ sampled in individuals of sex δ from the admixed population H at generation g. Thus,

Table 1 Recursion for the X-chromosomal and autosomal admixture fractions of a randomly chosen female or male from an admixed population at generation g, given a set of parents

Case	Female parent's population	Male parent's population	Probability	Admixture in males		Admixture in females	
				$H_{1,g,m}^A$	$H_{1,g,m}^X$	$H_{1,g,f}^A$	$H_{1,g,f}^X$
1	S_1	S_1	$s_{1,g-1}^f s_{1,g-1}^m$	1	1	1	1
2	S_1	H	$s_{1,g-1}^f h_{g-1}^m$	$\frac{1+H_{1,g-1,m}^A}{2}$	1	$\frac{1+H_{1,g-1,m}^A}{2}$	$\frac{1+H_{1,g-1,m}^X}{2}$
3	S_1	S_2	$s_{1,g-1}^f s_{2,g-1}^m$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$
4	H	S_1	$h_{g-1}^f s_{1,g-1}^m$	$\frac{1+H_{1,g-1,f}^A}{2}$	$H_{1,g-1,f}^X$	$\frac{1+H_{1,g-1,f}^A}{2}$	$\frac{1+H_{1,g-1,f}^X}{2}$
5	H	H	$h_{g-1}^f h_{g-1}^m$	$\frac{H_{1,g-1,f}^A+H_{1,g-1,m}^A}{2}$	$H_{1,g-1,f}^X$	$\frac{H_{1,g-1,f}^A+H_{1,g-1,m}^A}{2}$	$\frac{H_{1,g-1,f}^X+H_{1,g-1,m}^X}{2}$
6	H	S_2	$h_{g-1}^f s_{2,g-1}^m$	$\frac{H_{1,g-1,f}^A}{2}$	$H_{1,g-1,f}^X$	$\frac{H_{1,g-1,f}^A}{2}$	$\frac{H_{1,g-1,f}^X}{2}$
7	S_2	S_1	$s_{2,g-1}^f s_{1,g-1}^m$	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$
8	S_2	H	$s_{2,g-1}^f h_{g-1}^m$	$\frac{H_{1,g-1,m}^A}{2}$	0	$\frac{H_{1,g-1,m}^A}{2}$	$\frac{H_{1,g-1,m}^A}{2}$
9	S_2	S_2	$s_{2,g-1}^f s_{2,g-1}^m$	0	0	0	0

With two source populations and the contributions from the admixed population to itself, a random individual has nine possible parental pairings, for each of which the probability is listed.

$H_{1,g,\delta}^\gamma$ is the probability that a random site on an autosome ($\gamma = A$) or the X chromosome ($\gamma = X$) in a random individual of sex δ in population H in generation g ultimately traces to S_1. $H_{1,g,\delta}^\gamma$ tracks the admixture levels present after the entry of the admixture contributions involved in producing the admixed population at generation g.

In contrast to autosomal DNA, for which both female and male offspring receive a single copy of each chromosome from each parent, the X chromosome is inherited sex-specifically. That is, female offspring inherit one copy of the X chromosome from the mother and one from the father, but males inherit only a maternal copy. Therefore, whereas the autosomal admixture fractions sampled in females and males from the admixed population, $H_{1,g,f}^A$ and $H_{1,g,m}^A$, are identically distributed, the female and male X-chromosomal admixture fractions, $H_{1,g,f}^X$ and $H_{1,g,m}^X$, have different distributions. As a result, whereas Goldberg *et al.* (2014) could uncouple $H_{1,g,f}^A$ and $H_{1,g,m}^A$ and consider recursions for these two quantities separately (both were identical and only one needed to be studied), here we examine the distributions of $H_{1,g,f}^X$ and $H_{1,g,m}^X$ as a coupled pair of recursions.

Table 1 reports the probabilities that a random individual of a given sex from the admixed population has each possible set of parents, ℓ, as well as the admixture fraction of the individual for source population 1 conditional on the parental pairing. It provides a basis for recursively computing the fraction of admixture from S_1 for a random individual of sex δ from the admixed population at generation g, $H_{1,g,\delta}^X$. For $g = 1$, the founding event of the admixed population, the admixed population does not yet exist, so only cases 1, 3, 7, and 9, the four cases in which both parents are from the source populations, are considered.

Conditional on the previous generation, $H_{1,g,f}^X$ and $H_{1,g,m}^X$ are independent random variables. The female X-chromosomal admixture fraction depends on both the female and male admixture fractions in the previous generation, $H_{1,g-1,f}^X$ and $H_{1,g-1,m}^X$, but the male X-chromosomal admixture fraction depends only on $H_{1,g-1,f}^X$. Therefore, the female X-chromosomal admixture fraction can equivalently be written as a function of the female X-chromosomal admixture fractions in the previous generation, $H_{1,g-1,f}^X$, and from two generations ago, $H_{1,g-2,f}^X$.

Expectation of the X-Chromosomal Fraction of Admixture

Following Verdu and Rosenberg (2011) and Goldberg *et al.* (2014), we can find the expectation of the X-chromosomal admixture fraction of an individual of a specified sex randomly chosen from the admixed population. Using the law of total expectation to consider the random parental pairing L, we sum over the nine possible pairings ℓ (Table 1). For the mean X-chromosomal admixture fraction of a randomly chosen admixed individual of sex δ sampled at generation g, we have

$$\mathbb{E}\left[H_{1,g,\delta}^X\right] = \mathbb{E}_L\left[\mathbb{E}\left[H_{1,g,\delta}^X\middle|L\right]\right] = \sum_{\ell \in \left\{\begin{array}{c} S_1 S_1 \\ S_1 H \\ S_1 S_2 \\ H S_1 \\ H H \\ H S_2 \\ S_2 S_1 \\ S_2 H \\ S_2 S_2 \end{array}\right\}} P(L = \ell)\mathbb{E}\left[H_{1,g,\delta}^X\middle|L = \ell\right].$$

(2)

Applying Equation 2, using Table 1 to consider X chromosomes sampled in females and males from the admixed population, for the first generation, the initial condition is

$$\begin{bmatrix} \mathbb{E}\left[H_{1,1,\mathrm{f}}^{\mathrm{X}}\right] \\ \mathbb{E}\left[H_{1,1,\mathrm{m}}^{\mathrm{X}}\right] \end{bmatrix} = \begin{bmatrix} s_{1,0} \\ s_{1,0}^{\mathrm{f}} \end{bmatrix}. \tag{3}$$

For $g \geq 2$, we have

$$\begin{bmatrix} \mathbb{E}\left[H_{1,g,\mathrm{f}}^{\mathrm{X}}\right] \\ \mathbb{E}\left[H_{1,g,\mathrm{m}}^{\mathrm{X}}\right] \end{bmatrix} = \begin{bmatrix} s_{1,g-1} \\ s_{1,g-1}^{\mathrm{f}} \end{bmatrix} + \begin{bmatrix} \dfrac{h_{g-1}^{\mathrm{f}}}{2} & \dfrac{h_{g-1}^{\mathrm{m}}}{2} \\ h_{g-1}^{\mathrm{f}} & 0 \end{bmatrix} \begin{bmatrix} \mathbb{E}\left[H_{1,g-1,\mathrm{f}}^{\mathrm{X}}\right] \\ \mathbb{E}\left[H_{1,g-1,\mathrm{m}}^{\mathrm{X}}\right] \end{bmatrix}. \tag{4}$$

For the autosomes, for each sex δ, the expectation of the admixture fraction depends only on the total (non-sex-specific) contributions from the source populations and the corresponding expectation in the previous generation (Goldberg *et al.* 2014, equation 19),

$$\mathbb{E}\left[H_{1,g,\delta}^{\mathrm{A}}\right] = \begin{cases} s_{1,0} & g = 1 \\ s_{1,g-1} + h_{g-1}\mathbb{E}\left[H_{1,g-1,\delta}^{\mathrm{A}}\right] & g \geq 2. \end{cases} \tag{5}$$

By contrast, the expectations of the female and male X-chromosomal admixture fractions depend also on the sex-specific contributions in the previous generation (Equation 4).

Because the inheritance pattern is identical for X chromosomes and autosomes in females, the mean X-chromosomal admixture in females in Equation 4 matches that of the autosomes before equality of the female and male autosomal means is applied (Goldberg *et al.* 2014, equation 18). Whereas female and male autosomal admixture random variables are identically distributed and their expectations can be written with a generic δ, the corresponding random variables differ for X chromosomes, and the coupled recursion in Equations 3 and 4 cannot be quickly reduced to one equation. The mean X-chromosomal admixture in males (Equation 4), however, has a similar form to the autosomal mean (Equation 5).

We use the general mechanistic model presented here to derive closed-form expressions for the expected X-chromosomal admixture under two specific models of admixture, a single-admixture event and constant admixture over time.

Analyzing the Model in Special Cases

To analyze the properties of X-chromosomal admixture in relation to the female and male contributions of two source populations, we consider two specific cases of our general admixture model.

A single admixture event

First, we study a special case in which no further contributions from the source population occur after the admixed population is founded in generation $g = 1$. In this scenario of a single admixture event (Long 1991), $s_{1,g}^{\mathrm{f}} = s_{1,g}^{\mathrm{m}} = s_{2,g}^{\mathrm{f}} = s_{2,g}^{\mathrm{m}} = 0$, and $h_g^{\mathrm{f}} = h_g^{\mathrm{m}} = h_g = 1$, for all $g \geq 1$.

Applying Equations 3 and 4, we have for $g = 1$,

$$\begin{bmatrix} \mathbb{E}\left[H_{1,1,\mathrm{f}}^{\mathrm{X}}\right] \\ \mathbb{E}\left[H_{1,1,\mathrm{m}}^{\mathrm{X}}\right] \end{bmatrix} = \begin{bmatrix} s_{1,0} \\ s_{1,0}^{\mathrm{f}} \end{bmatrix}. \tag{6}$$

For $g \geq 2$,

$$\begin{bmatrix} \mathbb{E}\left[H_{1,g,\mathrm{f}}^{\mathrm{X}}\right] \\ \mathbb{E}\left[H_{1,g,\mathrm{m}}^{\mathrm{X}}\right] \end{bmatrix} = \begin{bmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbb{E}\left[H_{1,g-1,\mathrm{f}}^{\mathrm{X}}\right] \\ \mathbb{E}\left[H_{1,g-1,\mathrm{m}}^{\mathrm{X}}\right] \end{bmatrix}. \tag{7}$$

For $g \geq 2$, the expected admixture fraction for a male X chromosome is simply the expected admixture fraction for a female X chromosome from the previous generation, or $\mathbb{E}[H_{1,g,\mathrm{m}}^{\mathrm{X}}] = \mathbb{E}[H_{1,g-1,\mathrm{f}}^{\mathrm{X}}]$. We can use this identity between the expected female admixture in generation $g - 1$ and male admixture in generation g to simplify the system of equations. For $g \geq 3$, we then have

$$\mathbb{E}\left[H_{1,g,\mathrm{f}}^{\mathrm{X}}\right] = \frac{\mathbb{E}\left[H_{1,g-1,\mathrm{f}}^{\mathrm{X}}\right] + \mathbb{E}\left[H_{1,g-2,\mathrm{f}}^{\mathrm{X}}\right]}{2}. \tag{8}$$

In each generation, $\mathbb{E}[H_{1,g,\mathrm{f}}^{\mathrm{X}}]$ decreases by a factor of 2, suggesting that a closed-form solution might include a factor of 2^g. Therefore, denoting by y_g the quantity $2^{g-1}\mathbb{E}[H_{1,g,\mathrm{f}}^{\mathrm{X}}]$, using Equation 8, for $g \geq 3$, we have

$$\frac{y_g}{2^{g-1}} = \frac{y_{g-1}/2^{g-2} + y_{g-2}/2^{g-3}}{2}.$$

Multiplying both sides by 2^{g-1}, we have a recursion $y_g = y_{g-1} + 2y_{g-2}$, with $y_1 = s_{1,0}$ and $y_2 = s_{1,0} + s_{1,0}^{\mathrm{f}}$. Then for $g \geq 3$, y_g can be written $A_g s_{1,0} + B_g s_{1,0}^{\mathrm{f}}$, where A_g and B_g satisfy

$$A_g = A_{g-1} + 2A_{g-2} \tag{9}$$
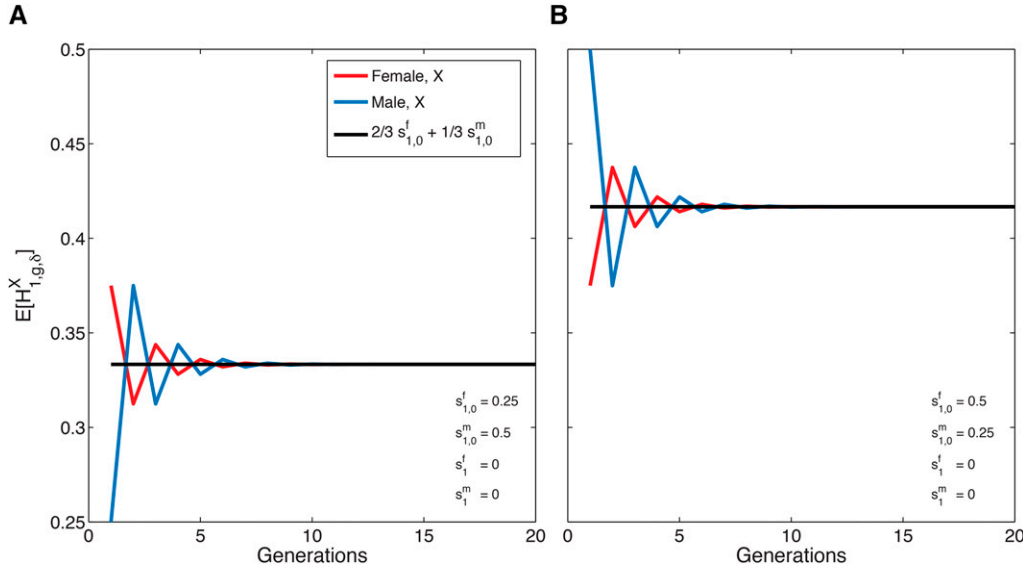
$$B_g = B_{g-1} + 2B_{g-2}, \tag{10}$$

with $A_1 = 1, A_2 = 1, B_1 = 0$, and $B_2 = 1$. Noting that $B_3 = 1$, we immediately observe that for $g \geq 3$, $B_g = A_{g-1}$, so that $y_g = A_g s_{1,0} + A_{g-1} s_{1,0}^{\mathrm{f}}$. Then

$$\mathbb{E}\left[H_{1,g,\mathrm{f}}^{\mathrm{X}}\right] = \frac{A_g s_{1,0} + A_{g-1} s_{1,0}^{\mathrm{f}}}{2^{g-1}}. \tag{11}$$

The recursion for A_g in Equation 9 with the initial conditions $A_0 = 0$ and $A_1 = 1$ gives the recursion for the Jacobsthal numbers, for which the closed form is $A_g = [2^g - (-1)^g]/3$ (Sloane 2015, entry A001045). We then have for $g \geq 1$,

$$\mathbb{E}\left[H_{1,g,\mathrm{f}}^{\mathrm{X}}\right] = \left[2 + \left(-\frac{1}{2}\right)^g\right] \frac{s_{1,0}^{\mathrm{f}}}{3} + \left[1 - \left(-\frac{1}{2}\right)^g\right] \frac{s_{1,0}^{\mathrm{m}}}{3}. \tag{12}$$

Figure 2 The mean of the X-chromosomal admixture fraction over time in females (red) and males (blue) from the admixed population, for a single admixture event (Equations 12 and 13). The mean X-chromosomal admixture (from source population 1) oscillates in approaching its limit (Equation 14). The limiting mean, shown in black, is the same for X chromosomes in both females and males. (A) $s_{1,0}^{\mathrm{f}} = 0.25$, $s_{1,0}^{\mathrm{m}} = 0.5$. (B) $s_{1,0}^{\mathrm{f}} = 0.5$, $s_{1,0}^{\mathrm{m}} = 0.25$.

Using Equation 7, for $g \geq 2$,

$$\mathbb{E}\left[H_{1,g,\mathrm{m}}^{\mathrm{X}}\right] = \left[2 + \left(-\frac{1}{2}\right)^{g-1}\right] \frac{s_{1,0}^{\mathrm{f}}}{3} + \left[1 - \left(-\frac{1}{2}\right)^{g-1}\right] \frac{s_{1,0}^{\mathrm{m}}}{3}. \tag{13}$$

For the case of a single admixture event, with no further gene flow from the source populations to the admixed population, we use Equations 12 and 13 to understand the behavior over time of the mean admixture on the X chromosome. Notably, unlike for the autosomes under a single-admixture scenario (Goldberg *et al.* 2014), the mean X-chromosomal admixture depends on the sex-specific contributions from the source populations, as well as on the time since admixture. With no sex bias ($s_{1,0}^{\mathrm{f}} = s_{1,0}^{\mathrm{m}} = s_{1,0}$), the mean X-chromosomal admixture is the same as the autosomal admixture, which is constant in time, equaling simply $s_{1,0}$, the total contribution from S_1.

Equations 12 and 13 oscillate over time (Figure 2), as they incorporate a negative fraction raised to a power. The mean X-chromosomal admixture does not immediately reach a long-term limit, as the difference in initial admixture values between females and males in the admixed population, followed by a transmission process in which females and males draw in different proportions from females and males in the previous generation, attenuates the homogenization of the female and male gene pools.

The long-term limit of the X-chromosomal admixture fraction is

$$\lim_{g \to \infty} \mathbb{E}\left[H_{1,g,\mathrm{f}}^{\mathrm{X}}\right] = \lim_{g \to \infty} \mathbb{E}\left[H_{1,g,\mathrm{m}}^{\mathrm{X}}\right] = \frac{2}{3}s_{1,0}^{\mathrm{f}} + \frac{1}{3}s_{1,0}^{\mathrm{m}}. \tag{14}$$

Because $\mathbb{E}[H_{1,g,\mathrm{m}}^{\mathrm{X}}] = \mathbb{E}[H_{1,g-1,\mathrm{f}}^{\mathrm{X}}]$ for all $g \geq 2$, the expected admixture fraction sampled in males approaches the same limit over time as that sampled in females. The limit of the mean X-chromosomal admixture fraction fits the 2:1 ratio by which

the X chromosome is inherited, following Equation 1, with the variables in Equation 1 viewed as the initial admixture values $s_{1,0}^{\mathrm{f}}$ and $s_{1,0}^{\mathrm{m}}$. Thus, the limiting admixture, but not the transient admixture, matches the simple linear combination.

For recent admixture, we can calculate the difference between the expected admixture under the model and the limiting value. For $g \geq 1$, we have

$$\mathbb{E}\left[H_{1,g,\mathrm{f}}^{\mathrm{X}}\right] - \lim_{g \to \infty} \mathbb{E}\left[H_{1,g,\mathrm{f}}^{\mathrm{X}}\right] = \frac{s_{1,0}^{\mathrm{f}} - s_{1,0}^{\mathrm{m}}}{3}\left(-\frac{1}{2}\right)^{g}, \tag{15}$$

and for $g \geq 2$,

$$\mathbb{E}\left[H_{1,g,\mathrm{m}}^{\mathrm{X}}\right] - \lim_{g \to \infty} \mathbb{E}\left[H_{1,g,\mathrm{m}}^{\mathrm{X}}\right] = \frac{s_{1,0}^{\mathrm{f}} - s_{1,0}^{\mathrm{m}}}{3}\left(-\frac{1}{2}\right)^{g-1}. \tag{16}$$

The differences in Equations 15 and 16 provide a measure of the difference of a transient single-admixture model from the simpler linear combination in Equation 1, which is constant in time and does not consider differences between the mean admixture in females and males from the admixed population. The $g \to \infty$ limit in Equation 14 agrees with Equation 1, but for small g, the differences in Equations 15 and 16 can be as large as $\pm 1/3$, decreasing by a factor of 2 each generation. For a fixed g, the maximal absolute difference occurs when the sex bias is largest, that is, when $s_{1,0}^{\mathrm{f}} - s_{1,0}^{\mathrm{m}}$ is 1 or -1. At the other extreme, with no sex bias and $s_{1,0}^{\mathrm{f}} = s_{1,0}^{\mathrm{m}}$, the simple linear combination exactly describes the X-chromosomal admixture throughout the history of the admixed population.

For six values of g, Figure 3 plots Equation 15 as a function of the difference between female and male contributions from S_1, $s_{1,0}^{\mathrm{f}} - s_{1,0}^{\mathrm{m}}$. With no sex bias, $s_{1,0}^{\mathrm{f}} = s_{1,0}^{\mathrm{m}}$, Equation 15 is zero, and our model follows Equation 1. Additionally, as g increases, the difference between our time-dependent model and Equation 1 becomes smaller, as Equation 1 gives the limiting behavior of the single-admixture model.
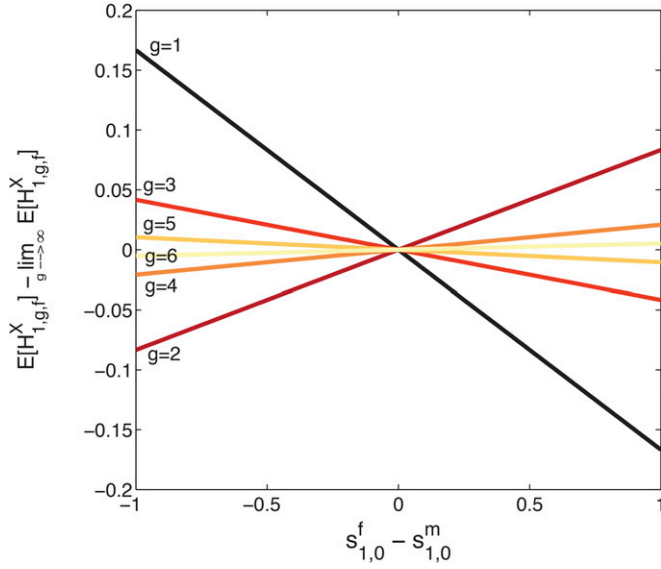
Figure 3 The difference between the expectation of the female X-chromosomal admixture fraction and its limit for $g \in [1, 6]$ as a function of the difference between female and male contributions from S_1 for a single admixture event (Equation 15). As $g \to \infty$, this quantity approaches zero. For small g, however, the limit $(2/3)s_{1,0}^f + (1/3)s_{1,0}^m$ (Equation 14) is a poor approximation. The difference oscillates so that the slope of the line is negative when g is odd and positive when g is even.

Constant, nonzero admixture over time

Next, we consider the special case of constant, nonzero contributions from the source populations to the admixed population over time. As in Goldberg *et al.* (2014), we can rewrite the sex-specific parameters as constants, $(s_{1,g}^f, s_{1,g}^m, s_{2,g}^f, s_{2,g}^m) = (s_1^f, s_1^m, s_2^f, s_2^m)$, for all $g \geq 1$. We maintain separate parameters for the founding contributions, $s_{1,0}^f$, $s_{1,0}^m$, $s_{2,0}^f$, and $s_{2,0}^m$. In this setting, h_f and h_m cannot both be 1, and at least one among s_1^f, s_1^m, s_2^f, and s_2^m must be nonzero.

Using a generating function approach, we derive a closed-form solution for the mean X-chromosomal admixture fractions, $\mathbb{E}[H_{1,g,f}^X]$ and $\mathbb{E}[H_{1,g,m}^X]$. The mean depends on the number of generations of constant admixture, g; the initial conditions $s_{1,0}^f$ and $s_{1,0}^m$; and the sex-specific contributions from the two source populations, s_1^f, s_1^m, s_2^f, and s_2^m. In the *Appendix*, we show that for each $g \geq 1$,

$$\mathbb{E}\left[H_{1,g,f}^X\right] = P_1 r_1^{g+1} + P_2 r_2^{g+1} + P_3, \tag{17}$$

where we have abbreviated a rather complex expression by using the quantities P_1, P_2, P_3, r_1, and r_2, functions that depend only on the admixture parameters and not on g. These quantities are defined in Equations A9–A13, with r_1 and r_2 <1 in absolute value. Using Equations A2 and 17, we can write an expression for the male X-chromosomal admixture fraction,

$$\mathbb{E}\left[H_{1,g,m}^X\right] = \begin{cases} s_{1,0}^f & g = 1 \\ s_1^f + h^f\left(P_1 r_1^g + P_2 r_2^g + P_3\right) & g \geq 2. \end{cases} \tag{18}$$

The limits of the mean X-chromosomal admixture fractions in Equations 17 and 18 are

$$\lim_{g \to \infty} \mathbb{E}\left[H_{1,g,f}^X\right] = \frac{s_1^f + s_1^m + s_1^f h^m}{s_1^f + s_1^m + s_2^f + s_2^m + h^m\left(s_1^f + s_2^f\right)} \tag{19}$$

$$\lim_{g \to \infty} \mathbb{E}\left[H_{1,g,m}^X\right] = \frac{2s_1^f + h^f s_1^m}{s_1^f + s_1^m + s_2^f + s_2^m + h^m\left(s_1^f + s_2^f\right)}. \tag{20}$$

These limits can also be obtained by an equilibrium approach. Denote $L_f = \lim_{g \to \infty} \mathbb{E}[H_{1,g,f}^X]$ and $L_m = \lim_{g \to \infty} \mathbb{E}[H_{1,g,m}^X]$. At equilibrium, an X chromosome in a male can ultimately trace to S_1 either through a mother from S_1 in the most recent generation or because it originates in an admixed female. Combining these terms, $L_m = s_1^f + h^f L_f$. An X chromosome in a female traces to S_1 either through the maternal X chromosome or through the paternal X chromosome: $L_f = (1/2)(s_1^f + h^f L_f) + (1/2)(s_1^m + h^m L_m)$. The solution to this pair of equations for L_f and L_m accords with Equations 19 and 20.

Unlike in the case of a single admixture event, the limit over time of the mean fraction of admixture from S_1 for a constant-admixture process is not a simple 2:1 weighting of the female and male contributions from S_1. Notably, recalling $s_1^f + h^f + s_2^f = s_1^m + h^m + s_2^m = 1$, the limiting admixture fraction from S_1 depends on the contributions both from S_1 and from S_2. The mean X-chromosomal admixture depends on the sex-specific contributions and cannot be reduced in terms of s_1 and s_2 only, as was possible for the autosomes (Goldberg *et al.* 2014, equation 37).

For the autosomes, the limiting mean is the ratio of contributions from S_1 to the total contributions from S_1 and S_2 combined (Goldberg *et al.* 2014, equation 37), $\mathbb{E}[H_{1,g,\delta}^A] = s_1/(s_1 + s_2)$. As was observed for the autosomes, the limiting mean X-chromosomal admixture in females can be viewed as the fractional contribution of X chromosomes from S_1 in relation to the total number of X chromosomes from both source populations. Therefore, in Equation 19, the numerator has the same form as the denominator, incorporating only contributions from S_1. X chromosomes from S_1 present in a population of females from the admixed population come from one of three origins in the limit, identified by the three terms of the numerator. They can be directly contributed from S_1 females or males, giving the terms s_1^f and s_1^m. Alternatively, they can be contributed from admixed males who in turn received them from S_1 females, producing the term $s_1^f h^m$; viewed in the limit, contributions from admixed females in the previous generation need not be considered, as an admixed female in the previous generation has the same equilibrium behavior as one in the current generation. Similar reasoning can be used to understand the limiting male admixture fraction, noting that the limiting mean for males is the sum of two quantities, the product of the limiting female mean and the fraction of females from the admixed population, h^f, and the new influx of female contributions from S_1, or s_1^f (Equation A2).
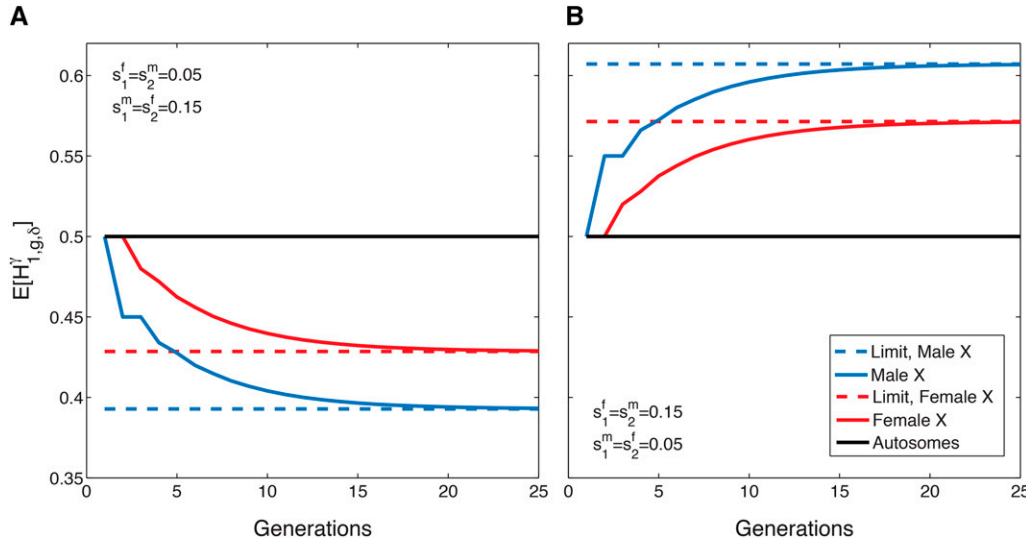
Figure 4 The expectation of the mean X-chromosomal and autosomal admixture fractions over time, with their associated limits, for constant ongoing admixture. (A) Male-biased admixture from population S_1. (B) Female-biased admixture from population S_1. The initial condition is ($s_{1,0}^f$, $s_{1,0}^m$, $s_{2,0}^f$, $s_{2,0}^m$) = $(0.5, 0.5, 0.5, 0.5)$. The autosomal admixture is constant over time because $s_{1,0} = s_{2,0} = 1/2$ and $s_1 = s_2$; it is the same in both A and B because it does not depend on the sex-specific contributions. The X-chromosomal admixture is different in females and males; it is smaller than the autosomal mean for male-biased admixture from population 1 and larger for female-biased admixture. The X-chromosomal mean is plotted using Equations 17–20. The autosomal mean uses equation 37 from Goldberg *et al.* (2014).

The limits over time of the means in Equations 19 and 20 depend only on the continuing contributions, not on the founding parameters. With no sex bias, $s_1^f = s_1^m = s_1$ and $h^f = h^m = h$, and we can simplify the limiting mean female and male X-chromosomal admixture fractions:

$$\lim_{g \to \infty} \mathbb{E}\left[H_{1,g,f}^X\right] = \lim_{g \to \infty} \mathbb{E}\left[H_{1,g,m}^X\right] = \frac{2s_1 + hs_1}{2 - h - h^2} = \frac{s_1}{1 - h}$$
$$= \frac{s_1}{s_1 + s_2}. \tag{21}$$

This limit is equivalent to the limiting mean of autosomal admixture from Verdu and Rosenberg (2011, equation 31). That is, with no sex bias, the limiting X-chromosomal mean matches the limiting autosomal mean, $\lim_{g \to \infty} \mathbb{E}\left[H_{1,g,\delta}^A\right] = \lim_{g \to \infty} \mathbb{E}\left[H_{1,g,f}^X\right] = \lim_{g \to \infty} \mathbb{E}\left[H_{1,g,m}^X\right]$.

Figure 4 plots the expectations of the X-chromosomal and autosomal admixture fractions over time for two scenarios with the same difference in female and male contributions, but different directions of sex bias. In Figure 4A, more males enter from S_1 and more females from S_2, leading to a lower mean admixture on the X chromosome than on the autosomes, with the mean in males smaller than the mean in females. Conversely, in Figure 4B, more females enter from S_1 and more males from S_2, leading to a larger mean for X chromosomes than for autosomes. The parameters are set so that for Figure 4, A and B, autosomal admixture is constant in time at $1/2$ (Goldberg *et al.* 2014, equation 37). Both panels illustrate a modified signal of the oscillation seen under a single admixture event, dampened by the ongoing contributions from the source populations. In both cases, the limiting mean X-chromosomal admixture in females lies intermediate

between the mean autosomal admixture and the mean X-chromosomal admixture in males.

Figure 5 plots the mean X-chromosomal admixture fraction for females and males (Equations 17 and 18) with the mean autosomal admixture fraction when both source populations have a female bias in contributions, $s_1^f > s_1^m$ and $s_2^f > s_2^m$. Although S_1 has an excess of females, the mean X-chromosomal admixture is less than the autosomal admixture. Whereas an excess of females from a given source population might be viewed as generating elevated X-chromosomal admixture, when both populations have a sex bias in the same direction, the source population with the larger sex bias dominates the signal. It is tempting to interpret lower X-chromosomal than autosomal admixture as a signal of male-biased admixture from S_1 and female-biased admixture from S_2, but Figure 5 plots an example where admixture is female biased in both source populations, and the simple interpretation of opposite biases in the two source populations is incorrect. A similar example could also have been produced with male bias in both source populations.

We note that in the limit as the parameters s_1^f, s_1^m, s_2^f, and s_2^m simultaneously approach zero, the constant-admixture model approaches the single-event model. Thus, we expect the limiting mean female and male X-chromosomal admixture fractions in the constant-admixture model to approach the corresponding limits in the single-event model. Indeed, by taking the limits of Equations 17 and 18 as s_1^f, s_1^m, s_2^f, and s_2^m approach 0, we obtain Equations 12 and 13, respectively. Interestingly, as we expect the trajectories of the X-chromosomal admixture fractions under constant admixture to approach corresponding single-event trajectories, we also expect the oscillatory pattern of the single-event model to occur in the constant-admixture model for
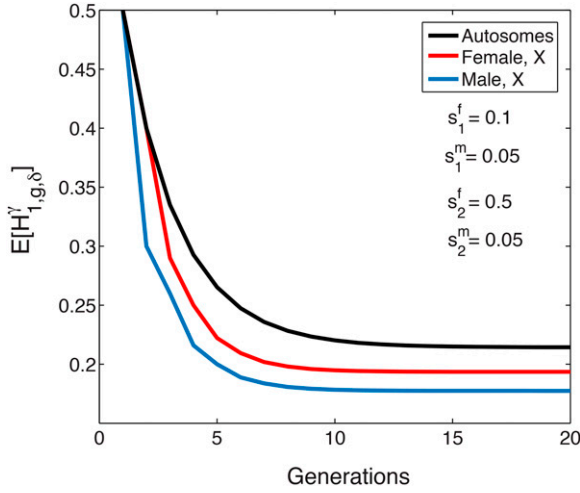
Figure 5 The expectation of the mean X-chromosomal and autosomal admixture fractions over time for constant admixture, with female-biased contributions from both source populations. The initial condition is $(s_{1,0}^f, s_{1,0}^m, s_{2,0}^f, s_{2,0}^m) = (0.5, 0.5, 0.5, 0.5)$. The mean X-chromosomal admixture from S_1, $\mathbb{E}[H_{1,g,\delta}^X]$, is smaller than the mean autosomal admixture, $\mathbb{E}[H_{1,g,\delta}^A]$, even though S_1 has an excess of females, $s_1^f > s_1^m$. The expectation of X-chromosomal admixture is plotted using Equations 17 and 18. The autosomal mean uses equation 37 from Goldberg *et al.* (2014).
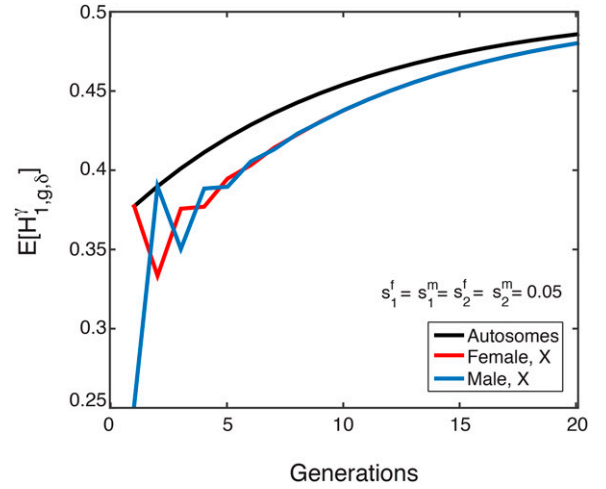
Figure 6 The expectation of the mean X-chromosomal and autosomal admixture fractions over time for constant admixture, in a case with small continuing contributions. The initial condition is $(s_{1,0}^f, s_{1,0}^m, s_{2,0}^f, s_{2,0}^m) = (0.25, 0.5, 0.75, 0.5)$. With sex bias in the founding generation and small continuing contributions, the mean X-chromosomal admixture has a pattern resembling the oscillating behavior seen for a single admixture event. The expectation of X-chromosomal admixture is plotted using Equations 17 and 18. The autosomal mean uses equation 37 from Goldberg *et al.* (2014).

sufficiently small s_1^f, s_1^m, s_2^f, and s_2^m. Indeed, we can find instances of oscillation around the limit in the constant-admixture model as well as in the single-admixture model, for example, for $s_1^f = s_1^m = s_2^f = s_2^m = 0.05$ and initial condition $(s_{1,0}^f, s_{1,0}^m, s_{2,0}^f, s_{2,0}^m) = (0.25, 0.5, 0.75, 0.5)$ (Figure 6). As the continuing contributions in this case have no sex bias, the mean female and male X-chromosomal admixture and the mean autosomal admixture tend to the same limit.

Sex-Biased Admixture in African-Americans

Studies of the genetic admixture history of African-American populations have consistently reported evidence for male-biased gene flow from Europe (Parra *et al.* 1998, 2001; Lind *et al.* 2007; Stefflova *et al.* 2009; Bryc *et al.* 2010a, 2015). Lind *et al.* (2007) and Bryc *et al.* (2015) estimated female and male contributions from Africa and Europe, using mean admixture estimates for the X chromosome and autosomes. Both analyses followed the linear combination method of Equation 1, in which the mean autosomal admixture averages the female and male contributions, while the mean X-chromosomal admixture weights the female and male contributions by 2/3 and 1/3, respectively. In our notation, the framework can be written

$$\mathbb{E}\left[H_{1,g,\delta}^A\right] = \frac{1}{2}s_1^f + \frac{1}{2}s_1^m \qquad (22)$$

$$\mathbb{E}\left[H_{1,g,\delta}^X\right] = \frac{2}{3}s_1^f + \frac{1}{3}s_1^m, \qquad (23)$$

with $\mathbb{E}[H_{1,g,\delta}^A]$ and $\mathbb{E}[H_{1,g,\delta}^X]$ estimated from data.

We have demonstrated, however, that in autosomal admixture models, Equation 22 suggests a single admixture event (Goldberg *et al.* 2014), and in X-chromosomal admixture models, Equation 23 suggests the limit of a single admixture process (replacing parameters for the continuing contributions in these equations by corresponding parameters for admixture in the initial generation, $s_{1,0}^f$ and $s_{1,0}^m$, as in Equation 14). For recent admixture in the single-admixture model, or in a model with continuing admixture after the founding of the admixed population, the mean admixture fractions from S_1 for the X chromosome depend on the sex-specific contributions in a more complex way, incorporating the sex-specific contributions from S_2 (Equations 12, 13, and 17–20). Use of Equations 22 and 23 implies consideration of a single-admixture model in its temporal limit or otherwise is unsuited to the single-admixture and continuing-admixture scenarios.

We use our refined predictions about $\mathbb{E}[H_{1,g,\delta}^X]$ to estimate the sex-specific contributions for an African-American population under two different models, a single-admixture event with various times since admixture and a constant-admixture process. Although both models underestimate the spatial and temporal complexity of the true admixture history of African-Americans, these approximations enable us to illustrate the way in which estimates of the sex bias depend on assumptions about the admixture model.

Genetic samples

Many admixture studies have reported X-chromosomal and autosomal admixture estimates for African-Americans (Table 2). For illustration, we focus on data from Cheng *et al.* (2009), who provided one of the largest samples, reporting

Table 2 Studies that published estimates of the X-chromosomal and autosomal admixture fractions of an African-American population.

	Mean ancestry estimates (%)				
	African autosome $\left(E\left[H^A_{1,g}\right]\right)$	European autosome $\left(E\left[H^A_{2,g}\right]\right)$	African X chromosome $\left(E\left[H^X_{1,g}\right]\right)$	European X chromosome $\left(E\left[H^X_{2,g}\right]\right)$	% female
Reference					
Cheng *et al.* (2009)	80.7	19.3	84.5	15.5	55.7
Bryc *et al.* (2015)	73.2	24.0	76.9	19.8	52.3
Bryc *et al.* (2010a)	77	23	81	19	66.3
Lind *et al.* (2007)	80	20	87.9	12.1	0
Tian *et al.* (2006)	81.8	18.2	87.1	12.9	Not reported

As the X chromosomes of females and males were considered jointly in the studies, we list the fraction of the sample that was female. In each study, the X chromosome shows elevated African ancestry compared to the autosomes. Values were compiled from quantities reported in the studies or by taking 1 minus an African estimate ("population 1") to obtain a European estimate ("population 2") or vice versa, with the following exceptions: For Bryc *et al.* (2015), the "autosome" estimates are genome-wide, including the X chromosome; the percentage of females was kindly provided by K. Bryc (personal communication). For Bryc *et al.* (2010a), we obtained an African X-chromosomal estimate by visual inspection of their figure 4g, and we counted the number of females from their supplementary material. For Tian *et al.* (2006), we computed the values from information reported in their figure 5 legend.

estimates of quantities close to those that appear in our model. Cheng *et al.* (2009) estimated admixture fractions of the X chromosome and autosomes for 15,280 African-Americans from 14 studies.

The mean admixture on the X chromosome reported by Cheng *et al.* (2009) combines females and males into a single estimate, whereas we consider separate quantities for X-chromosomal admixture in females and males. In our notation, the mean X-chromosomal admixture across all individuals, $\mathbb{E}[H^X_{1,g}]$, is a weighted average of admixture in females and males that takes into account the fraction of the sample that is female. Thus, for a sample divided into females and males in proportions p_f and $p_m = 1 - p_f$, respectively, we can write

$$\mathbb{E}\left[H^X_{1,g}\right] = p_f \mathbb{E}\left[H^X_{1,g,f}\right] + p_m \mathbb{E}\left[H^X_{1,g,m}\right]. \quad (24)$$

To compare admixture estimates from data to the mean admixture across individuals produced under mechanistic models, we compute the Euclidean distance between predictions for $\mathbb{E}[H^A_{1,g}]$, $\mathbb{E}[H^X_{1,g,f}]$, and $\mathbb{E}[H^X_{1,g,m}]$ and observed values in the data, weighting the predicted female and male X-chromosomal admixture by Equation 24. Denoting by q_A and q_X the observed mean admixture in a specific source group for autosomes and the X chromosome, we evaluate

$$D = \sqrt{\left[q_A - \mathbb{E}\left[H^A_{1,g}\right]\right]^2 + \left[q_X - \left(p_f \mathbb{E}\left[H^X_{1,g,f}\right] + p_m \mathbb{E}\left[H^X_{1,g,m}\right]\right)\right]^2}. \quad (25)$$

This quantity D measures the fit of a prediction under a model to the actual estimates of X-chromosomal and autosomal admixture for a data set that reports three quantities: estimates of admixture for a specific source population for autosomes and for the X chromosome and the fraction of the admixed sample consisting of females.

The exact timing of the onset of significant admixture in the African-American admixture process is unknown, but assuming a generation time of 20–27 years, the trans-Atlantic slave trade to North America can be regarded as having begun 14–20 generations ago (Eltis and Richardson 2010). As the mean X-chromosomal admixture fraction is generally near its limit for g in this range, both in the single-admixture model and in the constant-admixture model, the specific $g \in [14, 20]$ only minimally changes the results. For our constant-admixture analysis, we chose a single value $g = 15$ for consistency.

Single-admixture event

Using the mean X-chromosomal and autosomal admixture reported by Cheng *et al.* (2009) (Table 2), we calculated the sex-specific contributions $s^f_{1,0}$ and $s^m_{1,0}$ for specified values of $g \in [2, 5]$ and in the $g \to \infty$ limit. As Cheng *et al.* (2009) reported admixture in a combined sample of females and males, we used Equation 24 to calculate the X-chromosomal admixture fraction $\mathbb{E}[H^X_{1,g}]$ as a function of p_f, p_m, $\mathbb{E}[H^X_{1,g,f}]$, and $\mathbb{E}[H^X_{1,g,m}]$. Writing $\mathbb{E}[H^X_{1,g,f}]$ and $\mathbb{E}[H^X_{1,g,m}]$ as functions of g, $s^f_{1,0}$, and $s^m_{1,0}$ (Equations 12 and 13), we then generated an equation for $\mathbb{E}[H^X_{1,g}]$ in terms of p_f, p_m, g, $s^f_{1,0}$, and $s^m_{1,0}$, into which we inserted the reported values $q_X = 0.845$ for $\mathbb{E}[H^X_{1,g}]$ and $(p_f, p_m) = (0.557, 0.443)$. We obtained a second equation for the autosomal admixture fraction from Goldberg *et al.* (2014, equation 34), $\mathbb{E}[H^A_{1,g}] = s_{1,0} = (s^f_{1,0} + s^m_{1,0})/2$, using $q_A = 0.807$ for the empirical observation of $\mathbb{E}[H^A_{1,g}]$. We then solved the pair of equations for $s^f_{1,0}$ and $s^m_{1,0}$, with g fixed.

For small g (2, 3, 4, 5) as well as in the $g \to \infty$ limit, $s^f_{1,0}$, representing the contribution of females from the African source population, exceeds 0.9, and $s^f_{2,0}$, representing the female European contribution, is <0.1 (Table 3). For males, the African contribution $s^m_{1,0} \sim 0.7$, whereas the European contribution $s^m_{2,0} \sim 0.3$. Because of the oscillation of the mean X-chromosomal admixture (Figure 2), the estimated ratio of male to female contributions oscillates around the limiting values. For contributions from Europe, in the $g \to \infty$ limit we estimate $s^m_{2,0}/s^f_{2,0} \approx 3.89$. That is, in the temporal limit of a model of a single-admixture event, for each female from Europe, ~ 3.89 males contributed to the gene pool of African-Americans. The direction of this ratio is reversed for African

Table 3 Estimates of the sex-specific contributions from Africa and Europe to African-Americans, based on the Cheng *et al.* (2009) data, inferred under our model for a single-admixture event

Generations since admixture	African contributions			European contributions		
	$s_{1,0}^f$	$s_{1,0}^m$	$s_{1,0}^f/s_{1,0}^m$	$s_{2,0}^f$	$s_{2,0}^m$	$s_{2,0}^m/s_{2,0}^f$
2	0.943	0.671	1.40	0.057	0.329	5.77
3	0.912	0.702	1.30	0.088	0.298	3.39
4	0.926	0.688	1.35	0.074	0.312	4.22
5	0.919	0.695	1.32	0.081	0.305	3.77
∞	0.921	0.693	1.33	0.079	0.307	3.89

The estimates depend on the time since admixture. All estimates show male-biased contributions from Europe and female-based contributions from Africa.

contributions, with ∼1.33 females for every African male. If we assume that the admixture was recent, for example $g \leq 5$, then these values do vary substantially, with larger deviations from the limiting value of 3.89 occurring under more recent admixture.

Constant admixture over time

Assuming $g = 15$, we computed the mean female and male X-chromosomal admixture (Equations 17 and 18) and mean autosomal admixture given the four sex-specific contributions (Equation 5), on a grid of possible parameter values, $s_1^f, s_1^m, s_2^f, s_2^m \in [0, 1]$, using 0.01 increments. We fixed the initial values $s_{1,0}^f = s_{1,0}^m = 0.5$, although (except at small values of the continuing contributions) by $g = 15$ the mean nears its limit, erasing the signal of the initial conditions (Equations 19 and 20). We calculated the distance D between the Cheng *et al.* (2009) data and the computed values (Equation 25).

The parameters s_1^f, s_1^m, s_2^f, and s_2^m are not uniquely identifiable from the data, as we continue to have two equations, one describing autosomal admixture and one for X-chromosomal admixture, but now we consider four unknowns. We therefore examined sets of parameter values that produced $D \leq 0.01$. Distance cutoffs of 0.001 and 0.1 gave rise to similar ranges for each parameter. Thus, our analysis is based on the following algorithm:

1. Choose fixed values of g, $s_{1,0}^f$, and $s_{1,0}^m$. We fixed $g = 15$ and $s_{1,0}^f = s_{1,0}^m = 0.5$.
2. Choose $s_1^f, s_1^m, s_2^f, s_2^m \in [0, 1]$ on a four-dimensional grid, using 0.01 increments.
3. Plug the parameter values from steps 1 and 2 into Equations 5, 17, and 18 to calculate $\mathbb{E}[H_{1,g,\delta}^A]$, $\mathbb{E}[H_{1,g,f}^X]$, and $\mathbb{E}[H_{1,g,m}^X]$, respectively.
4. Plug the values for $\mathbb{E}[H_{1,g,\delta}^A]$, $\mathbb{E}[H_{1,g,f}^X]$, and $\mathbb{E}[H_{1,g,m}^X]$ from step 3 into Equation 25.
5. Choose a cutoff on D in Equation 25 that indicates a good fit of the parameters to the data. We use $D \leq 0.01$.

Figure 7 plots the set of parameter values that generate $D \leq 0.01$. Because the female and male contributions separately sum to one, the possible range of each set of contributions is represented by a unit simplex. Figure 7A plots the marginal distributions of the set of parameter values with $D \leq 0.01$. Con-

tributions from S_1 (Africans) vary over most of the permissible parameter space, but the S_2 contributions (Europeans) take their values from a narrower range. Figure 7B plots D on the space of female contributions, (s_1^f, h^f, s_2^f), for values of (s_1^m, h^m, s_2^m) fixed at the median of the set of parameter vectors that produce $D \leq 0.01$, $s_1^m = 0.47$, $h^m = 1 - s_1^m - s_2^m = 0.34$, and $s_2^m = 0.19$. Figure 7C plots D as a function of the male contributions, (s_1^m, h^m, s_2^m), with female contributions fixed at the median from the distribution of parameter values for which $D \leq 0.01$, $(s_1^f, h^f, s_2^f) = (0.64, 0.29, 0.07)$. In Figure 7, B and C, the parameter sets closest to the data in terms of D permit a large range of possible contributions from S_1, Africans, but only a smaller range of possible contributions, both female and male, from Europeans, S_2.

Figure 8A plots the natural logarithm of the ratio of male to female contributions for parameter sets with $D \leq 0.01$. With no sex bias, this quantity is zero; with more males than females, it is positive; and with more females than males, it is negative. Most points plotted from S_1, Africans, are female biased, whereas contributions from S_2, Europeans, are predominantly male biased. The median ratio of males to females from Europe is 2.67 males per female, compared to a median 1.32 females per male from Africa. Note, however, that these patterns do not hold for all parameter sets. For 26.15% of values with $D \leq 0.01$, more male than female contributions occur from Africa. For 9.25% of values with $D \leq 0.01$, more female than male contributions occur from Europe.

Although the pattern is initially surprising, scenarios of female bias from Europe or male bias from Africa accord with the phenomenon depicted in Figure 5. In the former scenario, both populations have an excess of females, and the strong female bias in the contributions from Africa can overwhelm the number of European X chromosomes in the gene pool of African-Americans, producing greater African admixture on the X chromosome than on autosomes. This effect is demonstrated in Figure 8B, which plots the logarithm of the ratio of male to female contributions from S_2 against the corresponding quantity from S_1. The ratios of male to female contributions in S_2 on the y-axis and S_1 on the x-axis are correlated. The values that are negative on the y-axis, indicating a female bias from S_2, also show the strongest female bias from S_1, that

is, the most negative $\ln(s_1^m/s_1^f)$ on the x-axis [an additional 2.34% of parameter sets show no sex bias in the European population; the female and male contributions are equal and $\ln(s_1^m/s_1^f) = 0$]. Analogously, the strongest male bias from Africa is also observed when the excess of males from Europe is largest (Figure 8B).

Discussion

Under a two-sex mechanistic admixture model with sex-biased admixture, we have demonstrated that the relationship between X-chromosomal and autosomal admixture fractions depends both on the time since admixture and on the model of admixture and does not simply follow a prediction from the fractions of X chromosomes and autosomes present in females. Using the mechanistic framework, we have reinterpreted African-American admixture values computed from a nonmechanistic perspective, estimating sex-specific parameters and levels of sex bias in African and European source populations. This analysis uncovers a counterintuitive case in which female-biased or male-biased contributions in the same direction occur both from Africans and from Europeans in a manner consistent with estimated mean X-chromosomal and autosomal admixture fractions in African-Americans, rather than an African female bias and a European male bias.

Estimating sex bias using X-chromosomal and autosomal admixture

Differences between X-chromosomal and autosomal admixture estimates are sometimes used to demonstrate the occurrence of sex bias, even without estimating sex-specific contributions (Wang *et al.* 2008; Bryc *et al.* 2010a,b; Cox *et al.* 2010; Beleza *et al.* 2012; Moreno-Estrada *et al.* 2013). Higher estimated admixture from a population S_1 for X chromosomes than for autosomes is taken as evidence of a female bias from S_1 and a male bias from a second population S_2. We have found, however, that the pattern is in principle compatible with two additional possibilities (Figure 8): a female bias from both populations, with a larger female bias from S_1, or a male bias from both populations, with a larger male bias from S_2.

We considered a general mechanistic model of admixture, focusing specifically on models of a single admixture event and constant admixture. For a single admixture event, mean autosomal admixture over time is constant, depending only on the total contributions from the source populations and not on the sex-specific contributions or time (Goldberg *et al.* 2014). Mean X-chromosomal admixture, on the other hand, is variable over time—in such a way that given the number of generations g and observed levels of X-chromosomal and autosomal admixture, the initial sex-specific contributions from the two source populations are identifiable. Because of the oscillation of the mean X-chromosomal admixture over time, depending on the time since admixture, the sex bias in admixture contributions can be either overestimated or

underestimated using the 2/3 and 1/3 linear combination (Equation 1).

For constant admixture, estimated mean X-chromosomal and autosomal admixture values do not uniquely identify the female and male contributions over time. Therefore, instead of point estimates for the sex-specific contributions and their ratio, we have reported a measure of the compatibility of parameter values and data over a range of values of the parameters (Figure 7 and Figure 8). It is possible, however, that by examining higher moments of the distribution of admixture estimates across individuals (Goldberg *et al.* 2014) or by simultaneously also considering the Y chromosome and mitochondrial DNA, the sex-specific contributions might become identifiable.

Theoretical population genetics of the X chromosome

The complex signature of admixture we have observed for the X chromosome is reminiscent of other X-chromosomal phenomena in theoretical population genetics, including results related to effective population size, allele-frequency dynamics, and numbers of ancestors. For effective population size, N_e, when female and male population sizes are equal, the ratio of X-chromosomal to autosomal values is 3/4 (Wright 1933; Caballero 1995; Hartl and Clark 2007). In the same way that various forms of sex difference between females and males—in such parameters as the number of individuals, the variance of reproductive success, and migration rates—cause the basic X-to-autosomal N_e ratio to differ from 3/4 (Caballero 1995; Charlesworth 2001; Ramachandran *et al.* 2004, 2008; Hammer *et al.* 2008; Keinan *et al.* 2009; Emery *et al.* 2010; Gottipati *et al.* 2011; Heyer *et al.* 2012; Arbiza *et al.* 2014), transient dynamics and ongoing admixture cause the X-chromosomal admixture fraction to differ from a linear combination of female and male contributions with coefficients 2/3 and 1/3.

The X-chromosomal admixture fraction in our single-admixture model also has similarities to the allele-frequency trajectory in the approach to Hardy–Weinberg equilibrium in a one-locus X-chromosomal model. In that model (Jennings 1916a; Wright 1969; Crow and Kimura 1970), the frequency of an allele on female X chromosomes depends on the frequency in both females and males in the previous generation, whereas the frequency on male X chromosomes depends only on the frequency in females. If an allele frequency differs between females and males at the first generation, then during the approach to equilibrium, the frequency in males matches the corresponding frequency in females of the previous generation; both the female and male frequencies oscillate around the same limit. The equilibrium frequency in turn has a 2/3 contribution from the initial female frequency and 1/3 from the male frequency.

All these phenomena from the one-locus model—female values dependent on both female and male values from the previous generation, male values dependent only on the female value, males lagging one generation behind females, oscillations around the same limit, and a limiting linear
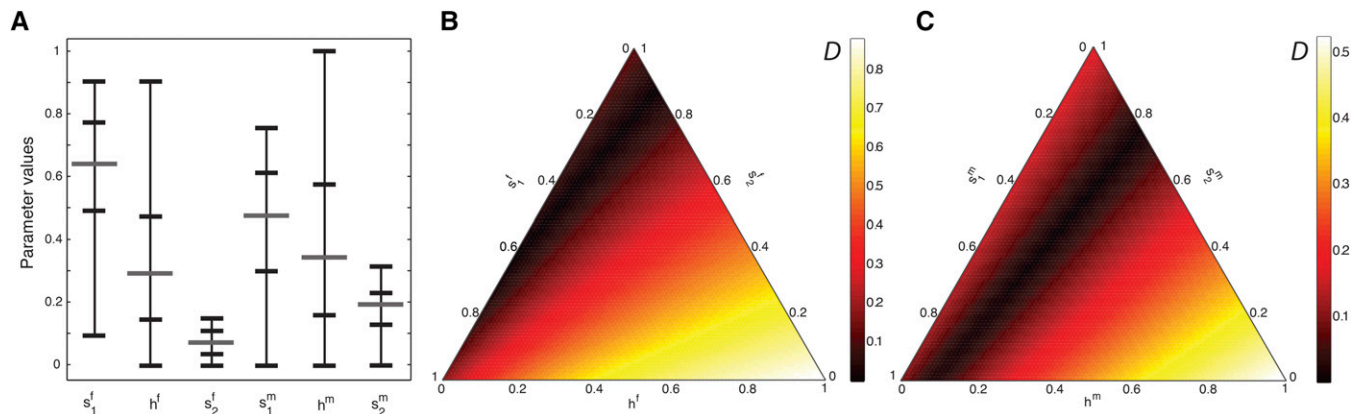
Figure 7 Sex-specific contributions estimated from the data of Cheng *et al.* (2009). (A) The range, median, and 25th and 75th percentiles of the sets of sex-specific contributions for which the Euclidean distance D (Equation 25) between model-predicted admixture and observed admixture from Cheng *et al.* (2009) was at most 0.01. The range of values for s_2^f and s_2^m, the contributions representing Europeans (S_2), is smaller than that representing Africans (S_1). (B) Female contributions as a function of D. (C) Male contributions as a function of D. For B, the male contributions are fixed at their median values producing $D \le 0.01$, $(s_1^m, h^m, s_2^m) = (0.47, 0.34, 0.19)$. For C, the female contributions are fixed in a corresponding manner, at $(s_1^f, h^f, s_2^f) = (0.64, 0.29, 0.07)$. Each side of the triangles in B and C represents one of three parameters that sum to 1, (s_1^m, h^m, s_2^m) and (s_1^f, h^f, s_2^f) in B and C, respectively. Parameter values were tested at increments of 0.01 for each quantity.

combination of 2/3 and 1/3—appear in our X-chromosomal admixture model with a single admixture event. An additional parallel lies in the difference between the X chromosome and the autosomes: In the one-locus model for the X chromosome, Hardy–Weinberg equilibrium is achieved in the limit, and for the autosomes, it is achieved in one generation. For admixture on the autosomes, the mean admixture is constant after the founding of the admixed population, and for the X chromosome, the mean X-chromosomal admixture approaches a limit rather than remaining constant in time.

Yet another connection to theoretical population genetics of the X chromosome comes from the form of the recursive sequence in Equation 11. The Jacobsthal sequence in the single-admixture model is obtained from a generalization of the Fibonacci sequence (Koshy 2001). Both the Jacobsthal and Fibonacci sequences appeared in breeding system models as early as a century ago (Jennings 1914, 1916a,b; Wright 1921), with the Jacobsthal sequence providing the numerators of allele frequencies at a sex-linked locus in a model that amounts mathematically to a special case of our single-event admixture model (Jennings 1916a,b); the Fibonacci sequence is well known as the number of genealogical ancestors of a haplodiploid system g generations ago (Basin 1963). Specifically, for the pair of X chromosomes in a female, the number of ancestors is entry $g + 2$ in the Fibonacci sequence F_n or, equivalently, the sum of the numbers of maternal and paternal ancestors, F_{g+1} and F_g, respectively. The number of genealogical ancestors for a male X chromosome in generation g is F_{g+1}. Our recursion $A_n = A_{n-1} + 2A_{n-2}$ (Equation 9) generates a different sequence, but its form is similar to the Fibonacci recursion $F_n = F_{n-1} + F_{n-2}$.

### Sex-biased admixture in African-Americans

Our analysis of sex-biased admixture in African-Americans has similarities but a number of noteworthy differences from earlier nonmechanistic analyses. Under our single-admixture model, the estimate of ~4 European males for every European female accords with estimated ratios of 3–4 from previous studies (Lind *et al.* 2007; Bryc *et al.* 2015). For even g, more African females and European males contribute than in the $g \to \infty$ limit; odd g values produce the opposite pattern.

For the constant-admixture model, the median estimated ratios of male to female contributions from Europe and Africa are ~2.67 males per female from Europe and ~1.32 females per male from Africa (Figure 7 and Figure 8). The estimated male bias in contributions from Europe is lower for a constant-admixture history than for a single-admixture event.

For both a single-admixture event and constant admixture over time, the point estimates of the ratio of males to females have a larger male contribution from Europe and female contribution from Africa. In the single-admixture model, the excess African ancestry on the X chromosome compared to the autosomes implies a female-biased contribution from Africa and a male-biased contribution from Europe. For constant admixture, surprisingly, a sex bias in one source alone can produce the observed pattern without a sex bias in the other (Figure 8). In fact, ~27% of the parameter sets most similar to the data ($D \le 0.01$) show no sex bias or have larger male than female contributions from Africa and ~12% show no sex bias or have larger female than male contributions from Europe. Note that in obtaining this result, we have given equal weight to all values of D below a cutoff, rather than giving more weight to parameter choices producing lower D; however, owing to the existence of ridges in the parameter space that have similarly small values of D, for different choices of the cutoff, similar results are produced.

An excess of African mitochondrial and European Y-chromosomal haplotypes has been described in African-Americans (Parra *et al.* 1998, 2001; McLean *et al.* 2005; Lind *et al.* 2007; Stefflova *et al.* 2009). Similar phenomena to those we observed for X chromosomes and autosomes could
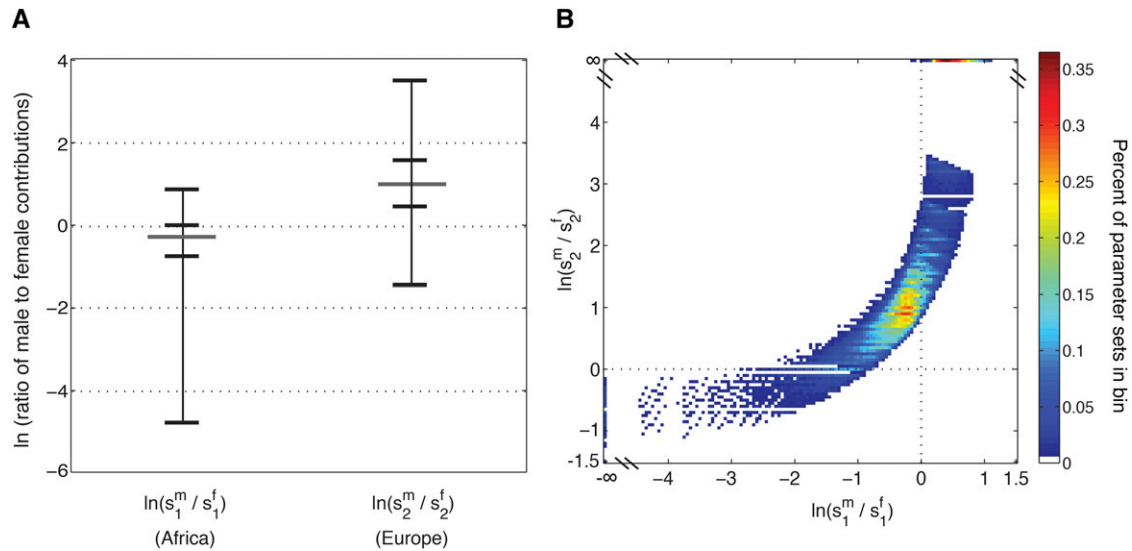
Figure 8 The natural logarithm of the ratio of male to female contributions in African-Americans, as inferred from the data of Cheng *et al.* (2009). (A) The range (excluding infinity, produced when a parameter value is zero), median, and 25th and 75th percentiles of the natural logarithm of the ratio of male to female contributions from S_1 (Africans) and S_2 (Europeans) separately for the sex-specific contributions that produced $D \leq 0.01$ (Equation 25). Values from Africans, S_1, are largely negative, or female biased, whereas contributions from Europeans, S_2, are mostly positive and male biased. Approximately 26.15% of the contributions from Africans are male biased and 9.25% from Europeans are female biased. This pattern is typically observed when a still larger sex bias occurs in the other population. (B) The logarithm of the ratio of male and female contributions from S_2 on the *y*-axis and the corresponding ratio for S_1 on the *x*-axis, plotted by the density of points in 0.05 square bins. For the cases with male bias in Africa ($\ln(s_1^m/s_1^f) > 0$), the level of male bias in Europe is also positive; for the cases with female bias in Europe ($\ln(s_2^m/s_2^f) < 0$), the level of female bias from Africa is also negative. Parameter sets in which at least one parameter is 0, and therefore we have values of $+\infty$ or $-\infty$ for $\ln(s_1^m/s_1^f)$ or $\ln(s_2^m/s_2^f)$, appear in bins on the edge of the plot for convenience. These bins contain a substantial number of parameter sets.

affect the Y–mitochondrial DNA (Y–mtDNA) comparisons: in other words, a strong excess of males from Europe compared to European females, even if more males than females contribute from African populations, could give rise to a larger fraction of European Y chromosomes in the African-American gene pool without a female bias for the African source population. Such a process could also explain the excess African ancestry of mtDNA.

Conclusions

We have presented a model to describe the effect of admixture on the X chromosome, deriving a theoretical framework that considers the impact of sex bias during the admixture process. The model can be used to understand the sex-specific contributions from source populations to an admixed population. We have found that because of model dependence, time dependence, and a lack of identifiability of admixture parameters from mean admixture alone, a variety of admixture processes and parameter values might be compatible with estimates of the mean admixture on X chromosomes and the autosomes. We have considered only ratios of female and male contributions rather than absolute numbers of female and male contributors, ignoring genetic drift. Both drift and other factors we have not studied—including mutation, recombination, and selection—can differentially affect the X chromosome and autosomes (Charlesworth 2001; Hedrick 2007; Pool and Nielsen 2007; Ellegren 2009; Labuda *et al.* 2010; Heyer *et al.* 2012; Ségurel *et al.* 2014; Veeramah *et al.* 2014),

potentially further complicating the estimation of admixture parameters. Examination in the context of sex-biased models of more detailed summaries of admixture patterns, including higher moments of the admixture distributions (Verdu and Rosenberg 2011; Gravel 2012; Goldberg *et al.* 2014) and distributions of the lengths of admixture tracts (Pool and Nielsen 2009; Gravel 2012; Liang and Nielsen 2014), will assist in refining the estimation of sex-biased admixture histories.

We demonstrated that female and male X-chromosomal admixtures have different expectations under mechanistic admixture models. With the exception of studies restricted to males (Lind *et al.* 2007), however, past studies of sex-biased admixture have generally reported composite X-chromosomal admixture estimates only in pooled collections of females and males. We recommend that such studies report estimates separately in females and males. Even if the values are similar, the difference between X-chromosomal admixture estimates in females and males contains information about parameters of the admixture process. Autosomal admixture estimates obtained only in females and only in males are identically distributed (Goldberg *et al.* 2014), so that a pooled estimate is more sensible than in the X-chromosomal case; nevertheless, autosomal admixture estimates reported separately in females and males can enable an informative comparison with a corresponding pair of X-chromosomal values.

Curiously, we have found that African-American admixture patterns do not necessarily imply a simultaneous African female bias and European male bias in the genetic ancestry

of the population and that in both source populations, it is in principle possible on the basis of genetic admixture patterns that both source populations had female bias or that both had male bias, albeit at different magnitudes. The latter interpretation, of male biases both in Europeans *and* in Africans is plausible in light of historical scholarship on demographic contributions to African-Americans (Fredrickson 1982; Painter 2006; Davis 2010; Eltis and Richardson 2010), documenting both the well-known contributions of male European slave owners and asymmetric mating practices involving African females and European males, as well as an overrepresentation of males among African slaves arriving in North America. We caution, however, that our focus has been on single-event and constant-admixture models, both of which approximate African-American admixture crudely, as neither one captures the time-varying nature of the contributions from the source groups. Exploration of additional scenarios, including models intermediate between these extreme cases, will be helpful for empirical problems. Nevertheless, both for African-Americans and for less well-documented cases, our model highlights that if a difference in X-chromosomal and autosomal admixture is observed, it is important to consider the possibility that rather than opposite sex biases in the two populations, both populations might have the same type of sex bias.

Acknowledgments

Literature Cited

Arbiza, L., S. Gottipati, A. Siepel, and A. Keinan, 2014 Contrasting X-linked and autosomal diversity across 14 human populations. Am. J. Hum. Genet. 94: 827–844.

Basin, S., 1963 The Fibonacci sequence as it appears in nature. Fibonacci Q 1: 53–56.

Bedoya, G., P. Montoya, J. García, I. Soto, S. Bourgeois *et al.*, 2006 Admixture dynamics in Hispanics: a shift in the nuclear genetic ancestry of a South American population isolate. Proc. Natl. Acad. Sci. USA 103: 7234–7239.

Beleza, S., J. Campos, J. Lopes, I. I. Araújo, A. H. Almada *et al.*, 2012 The admixture structure and genetic variation of the archipelago of Cape Verde and its implications for admixture mapping studies. PLoS ONE 7: e51103.

Berniell-Lee, G., S. Plaza, E. Bosch, F. Calafell, E. Jourdan *et al.*, 2008 Admixture and sexual bias in the population settlement of La Réunion Island (Indian Ocean). Am. J. Phys. Anthropol. 136: 100–107.

Bryc, K., A. Auton, M. R. Nelson, J. R. Oksenberg, S. L. Hauser *et al.*, 2010a Genome-wide patterns of population structure and admixture in West Africans and African Americans. Proc. Natl. Acad. Sci. USA 107: 786–791.

Bryc, K., C. Velez, T. Karafet, A. Moreno-Estrada, A. Reynolds *et al.*, 2010b Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. Proc. Natl. Acad. Sci. USA 107: 8954–8961.

Bryc, K., E. Y. Durand, J. M. Macpherson, D. Reich, and J. L. Mountain, 2015 The genetic ancestry of African Americans, Latinos, and European Americans across the United States. Am. J. Hum. Genet. 96: 37–53.

Bustamante, C. D., and S. Ramachandran, 2009 Evaluating signatures of sex-specific processes in the human genome. Nat. Genet. 41: 8–10.

Caballero, A., 1995 On the effective size of populations with separate sexes, with particular reference to sex-linked genes. Genetics 139: 1007–1011.

Casto, A. M., J. Z. Li, D. Absher, R. Myers, S. Ramachandran *et al.*, 2010 Characterization of X-linked SNP genotypic variation in globally distributed human populations. Genome Biol. 11: R10.

Charlesworth, B., 2001 The effect of life-history and mode of inheritance on neutral genetic variability. Genet. Res. 77: 153–166.

Cheng, C.-Y., W. H. L. Kao, N. Patterson, A. Tandon, C. A. Haiman *et al.*, 2009 Admixture mapping of 15,280 African Americans identifies obesity susceptibility loci on chromosomes 5 and X. PLoS Genet. 5: e1000490.

Cox, M. P., T. M. Karafet, J. S. Lansing, H. Sudoyo, and M. F. Hammer, 2010 Autosomal and X-linked single nucleotide polymorphisms reveal a steep Asian–Melanesian ancestry cline in eastern Indonesia and a sex bias in admixture rates. Proc. R. Soc. Lond. B Biol. Sci. 277: 1589–1596.

Crow, J. F., and M. Kimura, 1970 *An Introduction to Population Genetics Theory*. Burgess, Minneapolis.

Davis, F. J., 2010 *Who Is Black? One Nation's Definition*. Penn State Press, University Park, PA.

Ellegren, H., 2009 The different levels of genetic diversity in sex chromosomes and autosomes. Trends Genet. 25: 278–284.

Eltis, D., and D. Richardson, 2010 *Atlas of the Transatlantic Slave Trade*. Yale University Press, New Haven, CT.

Emery, L. S., J. Felsenstein, and J. M. Akey, 2010 Estimators of the human effective sex ratio detect sex biases on different timescales. Am. J. Hum. Genet. 87: 848–856.

Fredrickson, G. M., 1982 *White Supremacy: A Comparative Study of American and South African History*. Oxford University Press, Oxford.

Goldberg, A., P. Verdu, and N. A. Rosenberg, 2014 Autosomal admixture levels are informative about sex bias in admixed populations. Genetics 198: 1209–1229.

Gottipati, S., L. Arbiza, A. Siepel, A. G. Clark, and A. Keinan, 2011 Analyses of X-linked and autosomal genetic variation in population-scale whole genome sequencing. Nat. Genet. 43: 741–743.

Gravel, S., 2012 Population genetics models of local ancestry. Genetics 191: 607–619.

Hammer, M. F., F. L. Mendez, M. P. Cox, A. E. Woerner, and J. D. Wall, 2008 Sex-biased evolutionary forces shape genomic patterns of human diversity. PLoS Genet. 4: e1000202.

Hammer, M. F., A. E. Woerner, F. L. Mendez, J. C. Watkins, M. P. Cox *et al.*, 2010 The ratio of human X chromosome to autosome diversity is positively correlated with genetic distance from genes. Nat. Genet. 42: 830–831.

Hartl, D. L., and A. G. Clark, 2007 *Principles of Population Genetics*, Ed. 4. Sinauer Associates, Sunderland, MA.

Hedrick, P. W., 2007 Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. Evolution 61: 2750–2771.

Heyer, E., R. Chaix, S. Pavard, and F. Austerlitz, 2012 Sex-specific demographic behaviours that shape human genomic variation. Mol. Ecol. 21: 597–612.

Jennings, H. S., 1914 Formulae for the results of inbreeding. Am. Nat. 48: 693–696.

Jennings, H. S., 1916a   The numerical results of diverse systems of breeding. Genetics 1: 53–89.

Jennings, H. S., 1916b   The numerical results of diverse systems of breeding. Proc. Natl. Acad. Sci. USA 2: 45–50.

Keinan, A., J. C. Mullikin, N. Patterson, and D. Reich, 2009   Accelerated genetic drift on chromosome X during the human dispersal out of Africa. Nat. Genet. 41: 66–70.

Koshy, T., 2001   *Fibonacci and Lucas Numbers with Applications*. John Wiley & Sons, Hoboken, NJ.

Labuda, D., J.-F. Lefebvre, P. Nadeau, and M.-H. Roy-Gagnon, 2010   Female-to-male breeding ratio in modern humans—an analysis based on historical recombinations. Am. J. Hum. Genet. 86: 353–363.

Lambert, C. A., C. F. Connelly, J. Madeoy, R. Qiu, M. V. Olson *et al.*, 2010   Highly punctuated patterns of population structure on the X chromosome and implications for African evolutionary history. Am. J. Hum. Genet. 86: 34–44.

Liang, M., and R. Nielsen, 2014   The lengths of admixture tracts. Genetics 197: 953–967.

Lind, J. M., H. B. Hutcheson-Dilks, S. M. Williams, J. H. Moore, M. Essex *et al.*, 2007   Elevated male European and female African contributions to the genomes of African American individuals. Hum. Genet. 120: 713–722.

Long, J. C., 1991   The genetic structure of admixed populations. Genetics 127: 417–428.

Marcheco-Teruel, B., E. J. Parra, E. Fuentes-Smith, A. Salas, H. N. Buttenschøn *et al.*, 2014   Cuba: exploring the history of admixture and the genetic basis of pigmentation using autosomal and uniparental markers. PLoS Genet. 10: e1004488.

McLean, Jr., D. C., I. Spruill, G. Argyropoulos, G. P. Page, M. D. Shriver *et al.*, 2005   Mitochondrial DNA (mtDNA) haplotypes reveal maternal population genetic affinities of Sea Island Gullah-speaking African Americans. Am. J. Phys. Anthropol. 127: 427–438.

Moreno-Estrada, A., S. Gravel, F. Zakharia, J. L. McCauley, J. K. Byrnes *et al.*, 2013   Reconstructing the population genetic history of the Caribbean. PLoS Genet. 9: e1003925.

Painter, N. I., 2006   *Creating Black Americans: African-American History and Its Meanings, 1619 to the Present*. Oxford University Press, Oxford.

Parra, E. J., A. Marcini, J. Akey, J. Martinson, M. A. Batzer *et al.*, 1998   Estimating African American admixture proportions by use of population-specific alleles. Am. J. Hum. Genet. 63: 1839–1851.

Parra, E. J., R. A. Kittles, G. Argyropoulos, C. L. Pfaff, K. Hiester *et al.*, 2001   Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. Am. J. Phys. Anthropol. 114: 18–29.

Pool, J. E., and R. Nielsen, 2007   Population size changes reshape genomic patterns of diversity. Evolution 61: 3001–3006.

Pool, J. E., and R. Nielsen, 2009   Inference of historical changes in migration rate from the lengths of migrant tracts. Genetics 181: 711–719.

Quintana-Murci, L., C. Harmant, H. Quach, O. Balanovsky, V. Zaporozhchenko *et al.*, 2010   Strong maternal Khoisan contribution to the South African coloured population: a case of gender-biased admixture. Am. J. Hum. Genet. 86: 611–620.

Ramachandran, S., N. A. Rosenberg, L. A. Zhivotovsky, and M. W. Feldman, 2004   Robustness of the inference of human population structure: a comparison of X–chromosomal and autosomal microsatellites. Hum. Genomics 1: 87–97.

Ramachandran, S., N. A. Rosenberg, M. W. Feldman, and J. Wakeley, 2008   Population differentiation and migration: coalescence times in a two-sex island model for autosomal and X–linked loci. Theor. Popul. Biol. 74: 291–301.

Salzano, F. M., and M. Sans, 2014   Interethnic admixture and the evolution of Latin American populations. Genet. Mol. Biol. 37: 151–170.

Ségurel, L., M. J. Wyman, and M. Przeworski, 2014   Determinants of mutation rate variation in the human germline. Annu. Rev. Genomics Hum. Genet. 15: 47–70.

Sloane, N. J. A., 2015 The Online Encyclopedia of Integer Sequences. Available at: https://oeis.org/.

Stefflova, K., M. C. Dulik, A. A. Pai, A. H. Walker, C. M. Zeigler-Johnson *et al.*, 2009   Evaluation of group genetic ancestry of populations from Philadelphia and Dakar in the context of sex-biased admixture in the Americas. PLoS ONE 4: e7842.

Tian, C., D. A. Hinds, R. Shigeta, R. Kittles, D. G. Ballinger *et al.*, 2006   A genomewide single-nucleotide-polymorphism panel with high ancestry information for African American admixture mapping. Am. J. Hum. Genet. 79: 640–649.

Veeramah, K. R., R. N. Gutenkunst, A. E. Woerner, J. C. Watkins, and M. F. Hammer, 2014   Evidence for increased levels of positive and negative selection on the X chromosome versus autosomes in humans. Mol. Biol. Evol. 31: 2267–2282.

Verdu, P., and N. A. Rosenberg, 2011   A general mechanistic model for admixture histories of hybrid populations. Genetics 189: 1413–1426.

Verdu, P., N. S. Becker, A. Froment, M. Georges, V. Grugni *et al.*, 2013   Sociocultural behavior, sex-biased admixture, and effective population sizes in Central African Pygmies and non-Pygmies. Mol. Biol. Evol. 30: 918–937.

Wang, S., N. Ray, W. Rojas, M. V. Parra, G. Bedoya *et al.*, 2008   Geographic patterns of genome admixture in Latin American Mestizos. PLoS Genet. 4: e1000037.

Wilkins, J. F., 2006   Unraveling male and female histories from human genetic data. Curr. Opin. Genet. Dev. 16: 611–617.

Wilkins, J. F., and F. W. Marlowe, 2006   Sex-biased migration in humans: What should we expect from genetic data? BioEssays 28: 290–300.

Wright, S., 1921   Systems of mating II. The effects of inbreeding on the genetic composition of a population. Genetics 6: 124–143.

Wright, S., 1933   Inbreeding and homozygosis. Proc. Natl. Acad. Sci. USA 19: 411–420.

Wright, S., 1969   *The Theory of Gene Frequencies* (Evolution and the Genetics of Populations), Vol. 2. University of Chicago Press, Chicago.

Appendix: Solving for $\mathbb{E}[H_{1,g,\delta}^X]$ Under Constant Admixture

Here, we obtain the closed form for the mean X-chromosomal admixture fraction in a random female and a random male from the admixed population in the special case of constant admixture over time. Using the constant sex-specific parameters, s_1^f, s_1^m, s_2^f, and s_2^m, we rewrite the expectations in Equation 4. For $g = 1$, $\mathbb{E}\left[H_{1,1,f}^X\right] = s_{1,0}$, and for $g \geq 2$, we have

$$\mathbb{E}\left[H_{1,g,f}^X\right] = s_1 + \frac{1}{2}\left(h^f\mathbb{E}\left[H_{1,g-1,f}^X\right] + h^m\mathbb{E}\left[H_{1,g-1,m}^X\right]\right). \tag{A1}$$

Similarly, sampling males in the admixed population, for $g = 1$, $\mathbb{E}[H_{1,1,m}^X] = s_{1,0}^f$. For $g \geq 2$,

$$\mathbb{E}\left[H_{1,g,m}^X\right] = s_1^f + h^f\mathbb{E}\left[H_{1,g-1,f}^X\right]. \tag{A2}$$

Using Equations A1 and A2, we derive a generating function for the mean female X-chromosomal admixture fraction, which we then use to find closed-form expressions for $\mathbb{E}[H_{1,g,f}^X]$ and $\mathbb{E}[H_{1,g,m}^X]$. Because $\mathbb{E}[H_{1,g,m}^X]$ depends only on constants and $\mathbb{E}\left[H_{1,g-1,f}^X\right]$, we first find an expression for $\mathbb{E}[H_{1,g,f}^X]$, which we then use to report $\mathbb{E}[H_{1,g,m}^X]$.

First, using Equation A2, we can rewrite Equation A1 as a second-order recursion of a single variable,

$$\mathbb{E}\left[H_{1,g,f}^X\right] = \begin{cases} s_{1,0} & g = 1 \\ s_1 + \dfrac{1}{2}\left(s_{1,0}h^f + s_{1,0}^f h^m\right) & g = 2 \\ s_1 + \dfrac{h^f}{2}\mathbb{E}\left[H_{1,g-1,f}^X\right] + \dfrac{h^m}{2}\left(s_1^f + h^f\mathbb{E}\left[H_{1,g-2,f}^X\right]\right) & g \geq 3. \end{cases} \tag{A3}$$

We simplify the notation by defining $z_g = \mathbb{E}[H_{1,g,f}^X]$. For $g \geq 3$, we have

$$z_g = c_1 + c_2 z_{g-1} + c_3 z_{g-2}, \tag{A4}$$

with $c_1 = s_1 + s_1^f h^m/2$, $c_2 = h^f/2$, and $c_3 = h^f h^m/2$. Equation A3 gives $z_1 = s_{1,0}$ and $z_2 = s_1 + \frac{1}{2}(s_{1,0}h^f + s_{1,0}^f h^m)$.

Define a generating function $Z(x) = \sum_{g=1}^{\infty} z_g x^g$ whose coefficients z_g represent the values of $\mathbb{E}[H_{1,g,f}^X]$ in each generation. As the admixed population does not yet exist in generation 0, $\mathbb{E}[H_{1,g,f}^X]$ and $Z(x)$ are undefined for $g = 0$. For convenience, we therefore work with $W(x) = Z(x)/x$, setting $w_g = z_{g+1}$ for $g \geq 0$. We then have

$$W(x) = \sum_{g=0}^{\infty} z_{g+1}x^g = \sum_{g=0}^{\infty} w_g x^g, \tag{A5}$$

and $w_g = c_1 + c_2 w_{g-1} + c_3 w_{g-2}$ for $g \geq 2$. Using Equation A5, it follows that

$$\begin{aligned} W(x) &= w_0 + w_1 x + \sum_{g=2}^{\infty}(c_1 + c_2 w_{g-1} + c_3 w_{g-2})x^g \\ &= w_0 + w_1 x + c_1 \sum_{g=2}^{\infty} x^g + c_2 x \sum_{g=2}^{\infty} w_{g-1}x^{g-1} + c_3 x^2 \sum_{g=2}^{\infty} w_{g-2}x^{g-2} \\ &= w_0 + w_1 x + c_1\left(\frac{x^2}{1-x}\right) + c_2 x\left[W(x) - w_0\right] + c_3 x^2 W(x). \end{aligned} \tag{A6}$$

Solving for $W(x)$, we have

$$W(x) = \frac{(w_0 + w_1 x - c_2 w_0 x)(1-x) + c_1 x^2}{(1-x)(1 - c_2 x - c_3 x^2)}. \tag{A7}$$

We can decompose the expression in Equation A7, producing

$$W(x) = \frac{P_1}{1 - r_1 x} + \frac{P_2}{1 - r_2 x} + \frac{P_3}{1 - x}, \tag{A8}$$

where r_1 and r_2 are reciprocals of the two roots of $1 - c_2 x - c_3 x^2$,

$$r_1 = \frac{-c_2 + \sqrt{c_2^2 + 4c_3}}{2c_3} = \frac{h^{\mathrm{f}} + \sqrt{(h^{\mathrm{f}})^2 + 8h^{\mathrm{f}}h^{\mathrm{m}}}}{4} \tag{A9}$$

$$r_2 = \frac{-c_2 - \sqrt{c_2^2 + 4c_3}}{2c_3} = \frac{h^{\mathrm{f}} - \sqrt{(h^{\mathrm{f}})^2 + 8h^{\mathrm{f}}h^{\mathrm{m}}}}{4}. \tag{A10}$$

Setting Equation A7 equal to Equation A8, we have

$$P_1 = \frac{(w_0 r_1 + w_1 - c_2 w_0)(r_1 - 1) + c_1}{(r_1 - r_2)(r_1 - 1)} \tag{A11}$$

$$P_2 = \frac{(w_0 r_2 + w_1 - c_2 w_0)(r_2 - 1) + c_1}{(r_2 - r_1)(r_2 - 1)} \tag{A12}$$

$$P_3 = \frac{c_1}{(1 - r_1)(1 - r_2)}. \tag{A13}$$

The Taylor expansion of Equation A8 around $x = 0$ then gives

$$W(x) = P_1 \sum_{g=0}^{\infty} r_1^g x^g + P_2 \sum_{g=0}^{\infty} r_2^g x^g + P_3 \sum_{g=0}^{\infty} x^g,$$
$$= \sum_{g=0}^{\infty} (P_1 r_1^g + P_2 r_2^g + P_3) x^g.$$

Therefore, for $g \geq 0$, $w_g = P_1 r_1^g + P_2 r_2^g + P_3$, and the closed-form expression for the X-chromosomal female mean admixture fraction in generation $g \geq 1$, $\mathbb{E}[H_{1,g,\mathrm{f}}^{\mathrm{X}}]$, is $z_g = P_1 r_1^{g+1} + P_2 r_2^{g+1} + P_3$. We report this result in the main text as Equation 17, using it to obtain $\mathbb{E}[H_{1,g,\mathrm{m}}^{\mathrm{X}}]$ in Equation 18.

Because for $h^{\mathrm{f}}, h^{\mathrm{m}} \in [0, 1]$, r_1 monotonically increases in h^{f} and h^{m} and r_2 monotonically decreases, the maxima and minima of r_1 and r_2 occur at the boundaries of the closed interval $[0, 1]$. Using Equations A9 and A10, we have $r_1 \in [0, 1)$ and $r_2 \in (-1/2, 0]$; we exclude $r_1 = 1$ and $r_2 = -1/2$, as h^{f} and h^{m} cannot both be 1. Because $|r_1|, |r_2| < 1$, the mean X-chromosomal admixture fractions in Equations 17 and 18 approach limits as $g \to \infty$. Using Equations 17 and 18, we can find expressions for the limits of the mean of the X-chromosomal admixture fractions,

$$\lim_{g \to \infty} \mathbb{E}\left[H_{1,g,\mathrm{f}}^{\mathrm{X}}\right] = P_3$$

$$\lim_{g \to \infty} \mathbb{E}\left[H_{1,g,\mathrm{m}}^{\mathrm{X}}\right] = s_1^{\mathrm{f}} + h^{\mathrm{f}} P_3,$$

which can be simplified to give Equations 19 and 20.