

# A test for Hardy-Weinberg equilibrium on the X chromosome for sex-biased admixed populations

Daniel Backenroth<sup>1</sup> and Shai Carmi<sup>1</sup>

<sup>1</sup> Braun School of Public Health and community medicine, The Hebrew University of Jerusalem, Jerusalem, Israel

## Abstract

Genome-wide scans for deviations from Hardy-Weinberg equilibrium (HWE) are commonly applied to detect genotyping errors. In contrast to the autosomes, genotype frequencies on the X chromosome do not reach HWE within a single generation. Instead, if allele frequencies in males and females initially differ, they oscillate for a few generations towards equilibrium. Several populations world-wide have experienced recent sex-biased admixture, namely, their male and female founders differed in ancestry and thus in allele frequencies. Sex-biased admixture makes testing for HWE difficult on X, because deviations are *naturally* expected, even under random mating post-admixture and error-free genotyping. In this paper, we develop a likelihood ratio test and a  $\chi^2$  test that detect deviations from HWE on X while allowing for natural deviations due to sex-biased admixture. We demonstrate by simulations that our tests are powerful for detecting deviations due to non-random mating, while at the same time they do not reject the null under historical sex-biased admixture and random mating thereafter. We also demonstrate that when applied to 1000 Genomes project populations (e.g., as a quality control step), our tests reject fewer SNPs (among those showing frequency differences between the sexes) than other tests.

## Introduction

Testing for deviations from Hardy-Weinberg equilibrium (HWE) is an important quality control step in genome-wide association studies<sup>1-4</sup>. Extensive literature exists on HWE tests for the autosomes, from classic tests to recent work on Bayesian approaches, structured populations, sequenced or imputed genotypes, and software tools<sup>5-18</sup>. However, tests for HWE on the X chromosome have only been recently developed<sup>19-23</sup>. The importance of associations of X-linked variants with complex traits, particularly as a mechanism of sexual dimorphism, has been recently recognized<sup>24-32</sup>, and these developments underscore the importance of proper quality control on X, including testing for deviations from HWE.

A naive test for HWE on X would consider females only. However, such a test would implicitly assume an equal allele frequency between males and females. Indeed, a number of tests were recently proposed for *joint* testing of HWE in females as well as equality of allele frequencies between the sexes<sup>20-22</sup>. However, these tests ignore the possibility that allele frequencies in males and females would differ *naturally* due to sex-biased admixture.

While autosomal allele frequencies reach HWE within a single generation, it is well known that for X, in case male and female allele frequencies initially differ, perfect equilibrium is never reached<sup>33,34</sup>. The classical equations describing the evolution of allele frequencies on X, for an infinite population, are,

$$(1) \quad p_f(t+1) = \frac{p_m(t) + p_f(t)}{2},$$
$$(2) \quad p_m(t+1) = p_f(t),$$

where  $p_f(t)$  and  $p_m(t)$  are the male and female allele frequencies, respectively, at generation  $t$ . Starting with unequal allele frequencies at generation  $t = 0$ , the male and female frequencies oscillate while gradually stabilizing. Specifically<sup>34</sup>, if  $p_f(0) = 1$  and  $p_m(0) = 1$ , then  $p_f(t) = (2^{t+1} + (-1)^t)/(3 \cdot 2^n)$  and  $p_m(t) = (2^t - (-1)^t)/(3 \cdot 2^{t-1})$ . While equilibrium is approached exponentially quickly, if allele frequencies initially differ by a substantial amount, the frequency difference between the sexes can be non-negligible in the first few generations.

Recent sex-biased admixture has been known or identified for several populations, in particular in the Pacific and the Americas<sup>35-41</sup>. Moreover, admixture in these populations has often been cross-continental, which may have led to large initial frequency differences between the sexes. Thus, even if a population has been randomly mating since admixture, and even if SNPs are accurately genotyped, we may expect *natural* frequency differences to exist for some X-linked SNPs, along with natural deviations from HWE in females. Thus, it would be wrong to discard X SNPs due to HWE violation, in case the violation can be explained as a natural result of sex-biased admixture.

In this work, we developed a likelihood ratio test and a  $\chi^2$  test for HWE deviations on X, while permitting natural sex differences in frequency due to sex-biased admixture. This is achieved by taking into account the constraints imposed by Eqs. (1) and (2) on sex-specific frequency differences across generations. We show by simulations that our test has the expected size under the null, as well as power at least as high as existing tests for true deviations from the null (e.g., due to genotyping errors or inbreeding). Crucially, our test rejects HWE substantially less often compared to existing tests when HWE is violated due to historical sex-biased admixture in otherwise randomly mating populations. Finally, we show that in 1000 Genomes populations, our test rejects fewer SNPs among these for which frequency differences exist between the sexes.

## Methods

We denote the number of males and females in the sample as  $n_m$  and  $n_f$ , respectively, and the two alleles as A and B. The numbers of male A and B carriers are denoted  $m_A$  and  $m_B$ . The numbers of females with genotypes AA, AB, and BB are denoted  $f_{AA}$ ,  $f_{AB}$ , and  $f_{BB}$ . We denote by  $p_m$  and  $p_f$  the A allele frequencies in males and females, respectively.

We develop our likelihood ratio test based on the framework of You et al.<sup>21</sup>. These authors have defined the *inbreeding coefficient*  $\rho$  to represent deviations from HWE. Using  $\rho$ , the expected genotype frequencies in females can be written as

$$(3) \quad p_{AA} = (1 - \rho)p_f^2 + \rho p_f = p_f^2 + \rho p_f(1 - p_f),$$

$$(4) \quad p_{AB} = 2p_f(1 - p_f)(1 - \rho),$$

$$(5) \quad p_{BB} = (1 - \rho)(1 - p_f)^2 + \rho(1 - p_f) = (1 - p_f)^2 + \rho p_f(1 - p_f).$$

The null hypothesis of no deviations from HWE and no frequency difference between males and females is  $p_m = p_f = p$  and  $\rho = 0$ . We interpret here the parameter  $\rho$  more generally as a measure of the deviation from random mating in females, such that it can take any real value in  $[-1, 1]$ . (This guarantees

that all frequencies are in  $[0,1]$ .) The alternative hypothesis is  $p_m \neq p_f$  or  $\rho \neq 0$ . Denote the parameters of the model as  $\theta = (p_m, p_f, \rho)$ . The likelihood of observing of the data (genotype counts) is multinomial,

$$(6) L(\theta) = \binom{n_f}{f_{AA}, f_{AB}, f_{BB}} p_{AA}^{f_{AA}} p_{AB}^{f_{AB}} p_{BB}^{f_{BB}} \times \binom{n_m}{m_A} p_m^{m_A} (1 - p_m)^{m_B},$$

where  $p_{AA}$ ,  $p_{AB}$ , and  $p_{BB}$  are given by Eqs. (3), (4), and (5), respectively. You et al. have proposed an expectation-maximization algorithm to obtain the maximum likelihood estimates (MLE)  $\hat{\theta} = (\hat{p}_f, \hat{p}_m, \hat{\rho})$ .

Under the null hypothesis,  $p_m = p_f = p$  and  $\rho = 0$ , so  $\theta_0 = (p, p, 0)$ , and the likelihood reduces to

$$(7) L(\theta_0) = \binom{n_f}{f_{AA}, f_{AB}, f_{BB}} \binom{n_m}{m_A} 2^{f_{AB}} p^{2f_{AA} + f_{AB} + m_A} (1 - p)^{2f_{BB} + f_{AB} + m_B}.$$

Here, the MLE is trivial,  $\hat{\theta}_0 = (\hat{p}, \hat{p}, 0)$ , where  $\hat{p} = (2f_{AA} + f_{AB} + m_A)/(2n_f + n_m)$ . The likelihood ratio (LR) statistic is

$$(8) LR = 2 \ln \frac{L(\hat{\theta})}{L(\hat{\theta}_0)}.$$

The LR statistic is then asymptotically distributed (under the null) as a  $\chi^2$  distribution with two degrees of freedom, leading to a test we call the LRTP (likelihood ratio test for panmictic populations).

As explained above, the LRTP cannot accommodate “legitimate” frequency differences between the sexes due to sex-biased admixture. To address that, we reparametrize the model as follows. Instead of  $\theta = (p_f, p_m, \rho)$ , we write  $\theta = (p_{f,g}, p_{m,g}, \rho)$ , where  $p_{f,g}$  and  $p_{m,g}$  are the allele frequencies in females and males in the previous generation. With these parameters, the expected genotype frequencies in males in the current generation are

$$(9) p_{A,m,c} = p_{f,g}$$

$$(10) p_{B,m,c} = 1 - p_{f,g}.$$

This is analogous to Eq. (1), which is true because males receive X chromosomes only from females in the previous generation. In females, assume for the moment that once  $p_{f,g}$  and  $p_{m,g}$  are given, females in the current generation are the products of random mating. The expected genotype frequencies in females in the current generation would be

$$(11) p_{AA,f,c} = p_{m,g} p_{f,g},$$

$$(12) p_{AB,f,c} = p_{m,g} (1 - p_{f,g}) + p_{f,g} (1 - p_{m,g}),$$

$$(13) p_{BB,f,c} = (1 - p_{m,g}) (1 - p_{f,g}).$$

The above expressions reflect the fact that females receive one X chromosome from males and one from females. To incorporate deviations from random mating, we use again the parameter  $\rho$ . Analogously to the case of panmictic populations, we write the expected genotype frequencies in females in the current generation as

$$(14) p_{AA,f,c,\rho} = p_{m,g} p_{f,g} + \frac{\rho}{2} (p_{m,g} (1 - p_{f,g}) + p_{f,g} (1 - p_{m,g})),$$

$$(15) p_{AB,f,c,\rho} = (p_{m,g} (1 - p_{f,g}) + p_{f,g} (1 - p_{m,g})) (1 - \rho),$$

$$(16) p_{BB,f,c,\rho} = (1 - p_{m,g})(1 - p_{f,g}) + \frac{\rho}{2} (p_{m,g}(1 - p_{f,g}) + p_{f,g}(1 - p_{m,g})).$$

Note that the overall A allele frequency in females in the current generation is (for any  $\rho$ )

$$(17) p_{f,c} = p_{AA,f,c} + \frac{p_{AB,f,c}}{2} = \frac{p_{m,g} + p_{f,g}}{2},$$

as expected based on Eq. (2). Also note that here,  $\rho$  cannot take any value, as  $p_{AA,f,c,\rho}$ ,  $p_{AB,f,c,\rho}$ , and  $p_{BB,f,c,\rho}$  must all be within [0,1]. Our null hypothesis is that given the allele frequencies in the previous generation ( $p_{f,g}$  and  $p_{m,g}$ ), the genotypes of the current generation are determined by random mating, or  $\rho = 0$ . The alternative hypothesis is that there is a deviation from random mating, or  $\rho \neq 0$ . The likelihood of the data under the most general  $\theta$  is

$$(18) L(\theta) = \binom{n_f}{f_{AA}, f_{AB}, f_{BB}} p_{AA,f,c,\rho}^{f_{AA}} p_{AB,f,c,\rho}^{f_{AB}} p_{BB,f,c,\rho}^{f_{BB}} \times \binom{n_m}{m_A} p_{A,m,c}^{m_A} p_{B,m,c}^{m_B},$$

where  $p_{A,m,c}$ ,  $p_{B,m,c}$ ,  $p_{AA,f,c,\rho}$ ,  $p_{AB,f,c,\rho}$ , and  $p_{BB,f,c,\rho}$  are defined by Eqs. (9), (10), (14), (15), and (16), respectively. The MLE  $\hat{\theta} = (\hat{p}_{f,g}, \hat{p}_{m,g}, \hat{\rho})$  is obtained by taking the derivatives of (the logarithm of)  $L(\theta)$  and equating to zero. This results in a set of three equations, which are too tedious to reproduce here, and can be solved numerically to yield the MLE  $\hat{\theta} = (\hat{p}_{f,g}, \hat{p}_{m,g}, \hat{\rho})$ . In practice, we directly maximized the log-likelihood based on a grid search. (We discarded any parameter set  $\hat{\theta}$  leading to allele frequencies in the current generation outside the range [0,1] in Eqs. (14), (15), and (16).)

In the case of random mating,  $\rho = 0$ , and thus the parameters are  $\theta_0 = (p_{f,g}, p_{m,g}, 0)$ . The likelihood is

$$(19) L(\theta_0) = \binom{n_f}{f_{AA}, f_{AB}, f_{BB}} p_{AA,f,c}^{f_{AA}} p_{AB,f,c}^{f_{AB}} p_{BB,f,c}^{f_{BB}} \times \binom{n_m}{m_A} p_{A,m,c}^{m_A} p_{B,m,c}^{m_B},$$

where  $p_{AA,f,c}$ ,  $p_{AB,f,c}$ , and  $p_{BB,f,c}$  are defined by Eqs. (11), (12), and (13), respectively. Taking the derivatives of  $L(\theta_0)$  with respect to  $p_{f,g}$  and  $p_{m,g}$  and equating to zero results in the following pair of equations,

$$(20) \frac{f_{AA}}{\hat{p}_{m,g}} + \frac{f_{AB}(1-2\hat{p}_{f,g})}{(1-\hat{p}_{m,g})(1-\hat{p}_{f,g})} - \frac{f_{BB}}{1-\hat{p}_{m,g}} = 0$$

$$(21) \frac{f_{AA}}{\hat{p}_{f,g}} + \frac{f_{AB}(1-2\hat{p}_{m,g})}{(1-\hat{p}_{m,g})(1-\hat{p}_{f,g})} - \frac{f_{BB}}{1-\hat{p}_{f,g}} + \frac{m_A}{\hat{p}_{f,g}} - \frac{m_B}{1-\hat{p}_{f,g}} = 0$$

The solution of these equations yields the MLE under the null,  $\hat{\theta}_0 = (\hat{p}_{f,g}, \hat{p}_{m,g}, 0)$ . Here too, in practice we used a grid search to directly maximize the log-likelihood.

The likelihood ratio is then, as in Eq. (8),

$$(22) LR = 2 \ln \frac{L(\hat{\theta})}{L(\hat{\theta}_0)}.$$

Under the null, LR is asymptotically distributed as  $\chi^2$  with one degree of freedom, leading to a test we call LRTA (for admixture).

For comparison, we also consider the LRTG test of Graffelman and Weir<sup>22</sup>. In their test, the likelihood of the data is as in Eq. (6), except that  $p_{AA}$  and  $p_{AB}$  are parameters to be estimated, and  $p_{BB} = 1 - p_{AA} -$

$p_{AB}$  (i.e.,  $\theta = (p_{AA}, p_{AB}, p_m)$ ). The likelihood under the null is as in Eq. (7), and the likelihood ratio has a  $\chi^2$  distribution with two degrees of freedom.

Finally, we also use our results to propose a new  $\chi^2$  test<sup>22</sup>. Suppose we have used Eqs. (20) and (21) to obtain the MLE  $(\hat{p}_{f,g}, \hat{p}_{m,g})$ . The expected values for the genotypes of males and females under the null ( $\rho = 0$ ) are

$$(23) \text{ Expected } f_{AA}, f_{AB}, f_{BB}, m_A, m_B = n_f \hat{p}_{m,g} \hat{p}_{f,g}, n_f (\hat{p}_{m,g}(1 - \hat{p}_{f,g}) + \hat{p}_{f,g}(1 - \hat{p}_{m,g})), n_f (1 - \hat{p}_{m,g})(1 - \hat{p}_{f,g}), n_m \hat{p}_{f,g}, n_m (1 - \hat{p}_{f,g})$$

Then, given the observed values of  $f_{AA}, f_{AB}, f_{BB}, m_A, m_B$ , a standard  $\chi^2$  statistic can be calculated, which would be asymptotically distributed as  $\chi^2$  with one degree of freedom. We call this test  $\chi^2$ -ML.

We also note that instead of the MLE  $\hat{p}_{f,g}$  and  $\hat{p}_{m,g}$ , we could use a method of moments estimator, based on isolating  $p_f(t)$  and  $p_m(t)$  from Eqs. (1) and (2),

$$(24) \hat{p}_{f,g} = 2p_f - p_m = \frac{2(2f_{AA} + f_{AB})}{2n_f} - \frac{m_A}{n_m},$$

$$(25) \hat{p}_{m,g} = p_f = \frac{2f_{AA} + f_{AB}}{2n_f}.$$

These estimates can then be substituted in Eq. (23), and a  $\chi^2$  statistic can be calculated. We call this test  $\chi^2$ -MM. In practice, we found that the  $\chi^2$ -MM did not appropriately control the type I error rate (Table 1), and we did not report further experiments with that test.

## Results

We carried out several simulations to examine the behavior of our tests as compared to the LRTP and LRTG tests. We considered scenarios either under our tests' null hypothesis, as well as under a number of alternative hypotheses.

Our first simulation was designed to examine the tests under their null hypothesis, namely sex-biased admixture with random mating thereafter. We started with a population of 400 males and 400 females, and a single locus with an initial allele frequency of 80% in females and 30% in males. Given the allele frequencies in one generation, we calculated the expected genotype frequencies in the subsequent generation based on Eqs. (9)-(13). Then, the genotypes of 400 males and 400 females were drawn based on multinomial distributions having these expected frequencies. We repeated the process up to six generations after admixture, and repeated the simulation 1000 times.

In Table 1, we report the proportion of rejections (type I error rate) when running five tests on the above genotype counts: the LRTP test of You et al.<sup>21</sup> and the LRTG test of Graffelman and Weir<sup>22</sup>, both of which test for departures from either HWE in females or equality of allele frequencies between males and females; and the LRTA,  $\chi^2$ -ML, and  $\chi^2$ -MM tests we have developed here for sex-biased admixed populations (*Methods*). Our LRTA test and the  $\chi^2$ -ML test had an appropriate type I error rate (equal or close to the significance level  $\alpha = 0.05$ ), which is expected, because we simulated random mating post-admixture. In contrast, the LRTP and LRTG tests had much higher proportions of rejections, as expected due to the frequency differences between the sexes, which these tests are designed to detect. The type I

error rate decreased to its value under the null (0.05) after about  $\approx 6$  generations post-admixture, when allele frequency differences between males and females became very small. The  $\chi^2$ -MM test did not control the type I error rate as well as the LRTA test and the  $\chi^2$ -ML tests, possibly because the parameters (allele frequencies in the preceding generation) are not accurately estimated. We thus do not further consider this test.

Number of generations since admixture	L RTP	L RTG	L RTA (this paper)	$\chi^2$ -MM (this paper)	$\chi^2$ -ML (this paper)
1	1	1	<b>0.05</b>	0.22	<b>0.05</b>
2	0.97	0.97	<b>0.04</b>	0.09	<b>0.04</b>
3	0.47	0.48	<b>0.06</b>	0.07	<b>0.06</b>
4	0.17	0.18	<b>0.04</b>	0.05	0.05
5	0.09	0.09	<b>0.06</b>	<b>0.06</b>	<b>0.06</b>
6	0.10	0.10	<b>0.06</b>	<b>0.06</b>	<b>0.06</b>

Table 1. The proportion of rejections (Type I error rate) under random mating in a sex-biased admixed population. We compared the L RTP test (You et al. <sup>21</sup>), the L RTG test (Graffelman and Weir <sup>22</sup>), and the L RTA,  $\chi^2$ -MM, and  $\chi^2$ -ML tests developed in this paper. Our significance level was  $\alpha = 0.05$ . The lowest proportion in each row is highlighted in bold.

Our second simulation was designed to examine the power of the various tests under the alternative hypothesis of non-random mating. We considered one locus with an allele frequency of 80% in both males and females. We then calculated the expected genotype frequencies under one generation of mating, but this time with an inbreeding coefficient  $\rho$  equal to 0, 0.05, 0.1, 0.15, 0.2, 0.25, or 0.3, and simulated genotype frequencies in 400 females and 400 males based on the multinomial distribution with probabilities defined by Eqs. (9), (10), (14), (15), and (16). This simulation did not include sex-biased admixture, as the goal was to evaluate the power of our test under non-random mating, regardless of a history of admixture. We report the power of the various tests (at the 0.05 significance level and over 1000 repeats) in Table 2. The power of the  $\chi^2$ -ML test is always higher, followed closely by the L RTA test. The power of the L RTP and L RTG tests is slightly lower compared to our tests.

Inbreeding coefficient	L RTP	L RTG	L RTA (this paper)	$\chi^2$ -ML (this paper)
0	0.04	0.04	<b>0.05</b>	<b>0.05</b>
0.05	0.12	0.12	0.16	<b>0.19</b>
0.1	0.37	0.37	0.49	<b>0.52</b>
0.15	0.69	0.70	<b>0.82</b>	<b>0.82</b>
0.2	0.92	0.92	0.95	<b>0.97</b>

0.25	0.99	0.99	<b>1</b>	<b>1</b>
0.3	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>

Table 2. The proportion of rejections (power) of the various tests under non-random mating of increasing strengths (without admixture). The highest proportion in each row is highlighted in bold.

Our third simulation was designed to validate that the LRTA and  $\chi^2$ -ML tests are powerful also under sex-biased admixture. We used the same approach as in our first simulation (Table 1), i.e., sex-biased admixture followed by random mating, except that after one generation, non-random mating was assumed with an inbreeding coefficient equal to 0.1, 0.2, or 0.3. We report the power of the LRTA and  $\chi^2$ -ML tests (at the 0.05 significance level and over 50 repeats) in Table 3. The power of the tests is unaffected by the historical admixture event (cf. Table 2).

Inbreeding coefficient	LRTA (this paper)	$\chi^2$ -ML (this paper)
0.1	0.48	0.49
0.2	0.98	0.98
0.3	1	1

Table 3. The power of the LRTA and  $\chi^2$ -ML tests in populations with sex-biased admixture and increasing strength of non-random mating.

Finally, we applied our methods to real data from the 1000 Genomes project<sup>42</sup> (1kG). We selected American populations in which sex-biased, cross-continental recent admixture was likely. While admixture in these populations has mostly ended 5-10 generations ago (e.g.,<sup>43-46</sup>), some SNPs may have not yet reached equilibrium, or were affected by more recent minor gene flow events. Our goal in this analysis was to determine whether our tests indeed reject less SNPs due to deviation from HWE. However, as the power of our tests was higher compared to the other methods (Table 2), the proportion of rejected SNPs may not be informative, since many rejected SNPs could be genuinely affected by genotyping errors. We considered instead a subset of SNPs where there was a significant evidence for allele frequency difference between the sexes, based on the test of Zheng et al.<sup>19</sup>, at  $P < 0.05$ .

In Table 4, we report for each population the number of SNPs with a significant frequency difference between males and females, followed by the proportion of those SNPs rejected by each of the LRTP and LRTG tests as well as by our LRTA and  $\chi^2$ -ML tests. It can be seen that the proportion of rejected SNPs is lowest with our LRTA test. This result, along with the power simulations (Table 2), suggest that the LRTA test is likely to retain the maximal number of accurately genotyped SNPs for downstream analyses, while at the same time accurately detecting SNPs with true deviation from random mating. However, we note that in the absence of ground truth information on genotyping error status in 1kG, we cannot provide a formal proof of this claim.

Population	#sex-biased SNPs	LTRP	LRTG	LRTA (this paper)	$\chi^2$ -ML (this paper)
ACB (African Caribbeans in Barbados)	34206	0.38	0.38	<b>0.09</b>	0.12
ASW (Americans of African Ancestry in SW USA)	33217	0.25	0.25	<b>0.07</b>	0.12
CLM (Colombians from Medellin, Colombia)	22815	0.40	0.4	<b>0.11</b>	0.16
MXL (Mexican Ancestry from Los Angeles USA)	12778	0.40	0.4	<b>0.11</b>	0.16
PEL (Peruvians from Lima, Peru)	14562	0.37	0.37	<b>0.09</b>	0.10
PUR (Puerto Ricans from Puerto Rico)	21455	0.41	0.41	<b>0.11</b>	0.12

Table 4. The proportion of rejected SNPs (at  $\alpha = 0.05$ ) under the LTRP, LRTG, and LRTA tests in 1kG populations. We restricted our comparison to SNPs with a statistically significant frequency difference between the sexes. The lowest proportion in each row is highlighted in bold.

## Discussion

In this paper, we proposed new tests for deviations from HWE on the X chromosome for sex-biased admixed populations. The X chromosome is unique in that allele frequencies do not reach equilibrium within one generation after perturbation, even when the population is otherwise randomly mating and all genotypes are observed without errors. Thus, the X chromosome requires a specialized test for HWE, even beyond accounting for the different ploidy between the sexes. Here, we proposed new likelihood ratio and  $\chi^2$  tests to address this gap. We showed that our tests have the expected size (type I error rate) under sex-biased admixture and random mating thereafter, whereas other tests have high error rates, in particular when admixture was very recent. Additionally, our test has equal or higher power compared to the other tests considered. We also demonstrated that our tests reject fewer X chromosome SNPs in real 1000 Genomes populations. We thus recommend the application of our tests when performing quality control on the X chromosome. Our tests are available as an R package called HWadmiX at <https://github.com/dbackenroth/HWadmix>. Avenues for extending our approach can be the development of exact tests, Bayesian tests, or tests for multiple alleles.

## Acknowledgements

We thank Alon Keinan for discussions. S. C. thanks the German-Israeli Foundation for Scientific Research and Development (GIF) grant I-2489-407.6/2017 and the Israel Science Foundation (ISF) grant 407/17.

## Bibliography

1. Laurie, C.C. *et al.* Quality control and quality assurance in genotypic data for genome-wide association studies. *Genet Epidemiol* **34**, 591-602 (2010).



2. Anderson, C.A. *et al.* Data quality control in genetic case-control association studies. *Nat Protoc* **5**, 1564-73 (2010).
3. Turner, S. *et al.* Quality control procedures for genome-wide association studies. *Curr Protoc Hum Genet* **Chapter 1**, Unit1 19 (2011).
4. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-209 (2018).
5. Hernandez, J.L. & Weir, B.S. A disequilibrium coefficient approach to Hardy-Weinberg testing. *Biometrics* **45**, 53-70 (1989).
6. Ayres, K.L. & Balding, D.J. Measuring departures from Hardy-Weinberg: a Markov chain Monte Carlo method for estimating the inbreeding coefficient. *Heredity (Edinb)* **80 ( Pt 6)**, 769-77 (1998).
7. Emigh, T.H. A comparison of tests for Hardy-Weinberg equilibrium. *Biometrics* **36**, 627-42 (1980).
8. Bourgain, C., Abney, M., Schneider, D., Ober, C. & McPeck, M.S. Testing for Hardy-Weinberg equilibrium in samples with related individuals. *Genetics* **168**, 2349-61 (2004).
9. Wigginton, J.E., Cutler, D.J. & Abecasis, G.R. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* **76**, 887-93 (2005).
10. Rohlf, R.V. & Weir, B.S. Distributions of Hardy-Weinberg equilibrium test statistics. *Genetics* **180**, 1609-16 (2008).
11. Yu, C., Zhang, S., Zhou, C. & Sile, S. A likelihood ratio test of population Hardy-Weinberg equilibrium for case-control studies. *Genet Epidemiol* **33**, 275-80 (2009).
12. Wakefield, J. Bayesian methods for examining Hardy-Weinberg equilibrium. *Biometrics* **66**, 257-65 (2010).
13. Shriner, D. Approximate and exact tests of Hardy-Weinberg equilibrium using uncertain genotypes. *Genet Epidemiol* **35**, 632-7 (2011).
14. Graffelman, J., Sanchez, M., Cook, S. & Moreno, V. Statistical inference for Hardy-Weinberg proportions in the presence of missing genotype information. *PLoS One* **8**, e83316 (2013).
15. Graffelman, J., Jain, D. & Weir, B. A genome-wide study of Hardy-Weinberg equilibrium with next generation sequence data. *Hum Genet* **136**, 727-741 (2017).
16. Levene, H. On a Matching Problem Arising in Genetics. *Ann. Math. Stat.* **20**, 91 (1949).
17. Graffelman, J. Exploring Diallelic Genetic Markers: The HardyWeinberg Package. *J. Stat. Softw.* **64**, 1 (2015).
18. Hao, W. & Storey, J.D. Extending Tests of Hardy-Weinberg Equilibrium to Structured Populations. *BioRxiv* (2017).
19. Zheng, G., Joo, J., Zhang, C. & Geller, N.L. Testing association for markers on the X chromosome. *Genet Epidemiol* **31**, 834-43 (2007).
20. Puig, X., Ginebra, J. & Graffelman, J. A Bayesian test for Hardy-Weinberg equilibrium of biallelic X-chromosomal markers. *Heredity (Edinb)* **119**, 226-236 (2017).
21. You, X.P., Zou, Q.L., Li, J.L. & Zhou, J.Y. Likelihood Ratio Test for Excess Homozygosity at Marker Loci on X Chromosome. *PLoS One* **10**, e0145032 (2015).
22. Graffelman, J. & Weir, B.S. Testing for Hardy-Weinberg equilibrium at biallelic genetic markers on the X chromosome. *Heredity (Edinb)* **116**, 558-68 (2016).
23. Graffelman, J. & Weir, B.S. Multi-allelic exact tests for Hardy-Weinberg equilibrium that account for gender. *Mol Ecol Resour* **18**, 461-473 (2018).
24. Gao, F. *et al.* XWAS: A Software Toolset for Genetic Data Analysis and Association Studies of the X Chromosome. *J Hered* **106**, 666-71 (2015).
25. Chang, D. *et al.* Accounting for eXentricities: analysis of the X chromosome in GWAS reveals X-linked genes implicated in autoimmune diseases. *PLoS One* **9**, e113684 (2014).
26. Kukurba, K.R. *et al.* Impact of the X Chromosome and sex on regulatory variation. *Genome Res* **26**, 768-77 (2016).

27. Yap, C.X. *et al.* Dissection of genetic variation and evidence for pleiotropy in male pattern baldness. *Nat Commun* **9**, 5407 (2018).
28. Kudelka, M.R. *et al.* Cosmc is an X-linked inflammatory bowel disease risk gene that spatially regulates gut microbiota and contributes to sex-specific risk. *Proc Natl Acad Sci U S A* **113**, 14787-14792 (2016).
29. Li, Y.R. *et al.* Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. *Nat Med* **21**, 1018-27 (2015).
30. Traglia, M. *et al.* Genetic Mechanisms Leading to Sex Differences Across Common Diseases and Anthropometric Traits. *Genetics* **205**, 979-992 (2017).
31. Scelsi, M.A. *et al.* Genetic study of multimodal imaging Alzheimer's disease progression score implicates novel loci. *Brain* **141**, 2167-2180 (2018).
32. Khrantsova, E.A., Davis, L.K. & Stranger, B.E. The role of sex in the genomics of human complex traits. *Nat Rev Genet* (2018).
33. Jennings, H.S. The Numerical Results of Diverse Systems of Breeding. *Genetics* **1**, 53-89 (1916).
34. Rosenberg, N.A. Admixture Models and the Breeding Systems of H. S. Jennings: A GENETICS Connection. *Genetics* **202**, 9-13 (2016).
35. Bryc, K. *et al.* Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proc Natl Acad Sci U S A* **107 Suppl 2**, 8954-61 (2010).
36. Kim, S.K. *et al.* Population genetic structure and origins of Native Hawaiians in the multiethnic cohort study. *PLoS One* **7**, e47881 (2012).
37. Lie, B.A. *et al.* Molecular genetic studies of natives on Easter Island: evidence of an early European and Amerindian contribution to the Polynesian gene pool. *Tissue Antigens* **69**, 10-8 (2007).
38. Bonnen, P.E. *et al.* European admixture on the Micronesian island of Kosrae: lessons from complete genetic information. *Eur J Hum Genet* **18**, 309-16 (2010).
39. Lind, J.M. *et al.* Elevated male European and female African contributions to the genomes of African American individuals. *Hum Genet* **120**, 713-22 (2007).
40. Mathias, R.A. *et al.* A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome. *Nat Commun* **7**, 12522 (2016).
41. Jagadeesan, A. *et al.* Reconstructing an African haploid genome from the 18th century. *Nat Genet* **50**, 199-205 (2018).
42. Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).
43. Gravel, S. *et al.* Reconstructing Native American migrations from whole-genome and whole-exome data. *PLoS Genet* **9**, e1004023 (2013).
44. Gravel, S. Population genetics models of local ancestry. *Genetics* **191**, 607-19 (2012).
45. Moreno-Estrada, A. *et al.* Reconstructing the population genetic history of the Caribbean. *PLoS Genet* **9**, e1003925 (2013).
46. Baharian, S. *et al.* The Great Migration and African-American Genomic Diversity. *PLoS Genet* **12**, e1006059 (2016).