Testing for homogeneity of gametic disequilibrium across strata
Xiaolin Yin1, Wenqing Ma1, Manlai Tang2 and Jianhua Guo*1

Abstract

Background: Assessing the non-random associations of alleles at different loci, or gametic disequilibrium, can provide clues about aspects of population histories and mating behavior and can be useful in locating disease genes. For gametic data which are available from several strata with different allele probabilities, it is necessary to verify that the strata are homogeneous in terms of gametic disequilibrium.

Results: Using the likelihood score theory generalized to nuisance parameters we derive a score test for homogeneity of gametic disequilibrium across several independent populations. Simulation results demonstrate that the empirical type I error rates of our score homogeneity test perform satisfactorily in the sense that they are close to the pre-chosen 0.05 nominal level. The associated power and sample size formulae are derived. We illustrate our test with a data set from a study of the cystic fibrosis transmembrane conductance regulator gene.

Conclusion: We propose a large-sample homogeneity test on gametic disequilibrium across several independent populations based on the likelihood score theory generalized to nuisance parameters. Our simulation results show that our test is more reliable than the traditional test based on the Fisher’s test of homogeneity among correlation coefficients.

Background

Measuring gametic disequilibrium can provide important information about aspects of population histories and mating behavior [1] and can be useful in locating disease genes [2]. The term gametic disequilibrium is used in this article instead of the traditional term linkage disequilibrium to measure the extent of non-random association between unlinked loci [3]. Various measures of gametic disequilibrium have been proposed [4-6], ranging from pairs of diallelic loci model to multiple multiallelic loci model. In this article, we consider the gametic disequilibrium which is defined as the difference between the gametic probability and its expected probability under the assumption of no statistical association of alleles, and the gametic disequilibrium calculations are based on two-allele, two-locus model [7].

Consider two loci, A and B, each having two possible alleles (A0, A1) and (B0, B1), respectively. With two loci and two alleles, there are four possible gametes, namely, A0B0, A0B1, A1B0 and A1B1. The gametic disequilibrium between the two loci is defined by
\[
D = p_{A_iB_i} - p_{A_i}p_{B_i},
\]

where \( p_{A_i} \) and \( p_{B_i} \) denote the allele probabilities of \( A_i \) and \( B_i \), respectively. Here, \( p_{A_iB_i} \) denotes the gametic probability of \( A_iB_i \), \( i, j = 0, 1 \). Suppose that the gametic data are available from \( K \) strata and let \( p_{ijk} \) denote the gametic probability of array of \( A_iB_i \) for the \( k \)-th stratum, \( i, j = 0, 1; k = 1, ..., K \). \( \sum_k p_{ijk} = 1 \) for each \( k \). According to the relationship between allelic probability and gametic probability, the allele probabilities of \( A_0, A_1, B_0 \) and \( B_1 \) are derived as \( p_{00i} \), \( p_{11i} \), \( p_{01i} \) and \( p_{10i} \) respectively. Here "+" denote the summation over 0 and 1, for example, \( p_{0+i} = p_{00k} + p_{01k} \). For stratum \( k (k = 1, ..., K) \), the gametic disequilibrium is calculated as

\[
D_k = p_{11k} - p_{10k}p_{1+k}. \]

It is easy to show that \( D_k \) is bounded by

\[
D_k,\min \leq D_k \leq D_k,\max,
\]

where \( D_k,\min = \min \{ p_{11i} + p_{10i}p_{1+k}, p_{01i} + p_{01i}p_{1+k} \} \), \( D_k,\max = \min \{ p_{11i}p_{01k}, p_{10i}p_{01k} \} \). Testing for the homogeneity of gametic disequilibrium among strata can be informative in differentiating among the evolutionary agents generating them in natural population [8]. Detecting gametic disequilibrium can be informative in mapping gene and providing meaningful clues of population evolution. Combining the evidence of gametic disequilibrium across several strata may be more sufficient to support the clues, in contrast to analysis with each strata. In this case, it is crucial to test the homogeneity of gametic disequilibrium across strata before combining the data. For this purpose, it is interesting to consider the following hypothesis

\[
H_0 : D_1 \cup D_K \quad \text{versus} \quad H_1 : D_i \neq D_j \text{for at least a pair } i \neq j. \tag{1}
\]

Weir [9] recommended a homogeneity test on gametic disequilibrium, based on Fisher’s test of homogeneity among correlation coefficients [10]. In his method, the gametic disequilibrium \( D_k \) is first transformed to a correlation coefficient \( r_k \) by

\[
r_k = D_k / \sqrt{p_{11k}p_{01k}p_{1+k}p_{0+k}}, \quad r_k \text{ is then transformed to a normal variable } z_k \text{ by Fisher’s } z \text{ transformation, and a weighted sum of squares of the } z \text{ values which has } \chi^2 \text{ distribution with } K - 1 \text{ degrees of freedom is finally proposed for testing homogeneity of gametic disequilibrium. As pointed out by Zapata and Alvarez [8], this test is actually for homogeneity of } r \text{ values instead of } D \text{ values. They may not be equivalent when the allele probabilities are different across strata. Instead, Zapata and Alvarez [8] suggested the use of the normalized difference } D' \tag{11}. Specifically, } D_k \text{ is the ratio of } D_k \text{ to } D_{k,\max} \text{ when } D_k > 0, \text{ or the ratio of } D_k \text{ to } -D_{k,\min} \text{ when } D_k < 0. \text{ Zapata and Alvarez obtained the bias-corrected confidence interval for each } D' \text{ value across strata via the bootstrap method. Hence, acceptance or rejection of homogeneity of } D' \text{ values can be determined by evaluating the obtained confidence intervals. For the example considered in Zapata and Alvarez [8], there is no intersection for the confidence intervals obtained from all strata. Hence, one has evidence to reject the null hypothesis of homogeneity. Unfortunately, Zapata and Alvarez [8] did not discuss the decision rules for cases such as intersections exist but the extent are different. Hence, no rigorous rule based on this confidence interval approach was proposed and this makes their method less practicable. However, no rigorous rule based on this confidence interval approach was proposed and this makes their method less practicable. It should be noted that the homogeneity test of either } r \text{ values or } D' \text{ values is not equivalent to the homogeneity test of } D \text{ values. In particular, transformation } D' \text{ only guarantees that the range of } D' \text{ is } [-1, 1]. \text{ However, there remains difficulties in interpreting the value of } D'. \text{ Lewontin [11] noted that values of } D' \text{ at different loci and in different populations tend to vary with the values of the allele probabilities, so that the problem of cross-locus and cross-population comparisons is not fully overcome by the use of } D'. \text{ In this article, without doing any transformation, we develop an asymptotic homogeneity test directly based on } D \text{ values via score method.}

Methods

Homogeneity test

Let \( x_{ijk} (i, j = 0, 1 \text{ and } k = 1, \cup, K) \) be the number of the gamete \( A_iB_i \) in the \( k \)-th stratum with the total gametes being \( n_k = x_{00k} + x_{01k} + x_{10k} + x_{11k} \). Let \( M(n_{ij} \{ p_{ijk} \}) \) denote the quadrinomial distribution with parameter vector \( (p_{00k}, p_{01k}, p_{10k}, p_{11k}) \). Thus, we have \( \{x_{ijk} : i, j = 0, 1 \} \sim M(n_{ij} \{ p_{ijk} \}) \) for \( k = 1, ..., K \). The homogeneity hypothesis in (1) is of interest in this article. Here, we assume that \( K \) is fixed and \( n_k \) is sufficiently large for \( k = 1, 2, ..., K \). Noticing that \( p_{00k} = p_{00i}p_{0+k} + p_{01i}p_{1+k} - p_{11i}p_{0+k} - D_k \), \( p_{11k} = p_{11i}p_{0+k} + D_k \), the log-likelihood for the \( k \)-th stratum can be expressed in terms of \( D_k \) and \( p_{1+k} \) and \( p_{1+k} (k = 1, ..., K) \).

\[
l_k(D_k, p_{1+k}, p_{1+k}) = \ln(p_{00k}p_{0+} + D_k) + x_{01k}\ln(p_{01k} - D_k) + x_{10k}\ln(p_{10k} - D_k) + x_{11k}\ln(p_{11k} + D_k),
\]
where \( p_{0+k} = 1 - p_{1+k} \), \( p_{0+k}^* = 1 - p_{1+k}^* \). Let \( D \) denote the common gametic disequilibrium under \( H_0 \), \( \hat{p}_{1+k} = (p_{1+k}, ..., p_{1+k})' \), and \( \hat{p}_{1+k}^* = (p_{1+k}^*, ..., p_{1+k}^*)' \) denote the nuisance parameter vectors. Under \( H_0 \), the total log-likelihood for all \( K \) strata is given by

\[
l(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*) = \sum_{k=1}^{K} l_k(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*).
\]

Hence, the efficient scores for the \( k \)-th stratum (i.e., the first order derivatives of \( l_k \)) with respect to \( D \), \( \hat{p}_{1+k} \), and \( \hat{p}_{1+k}^* \) are given by

\[
S_{\partial l/\partial D}(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*) = \frac{\partial l_k(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*)}{\partial D},
S_{\partial l/\partial \hat{p}_{1+k}}(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*) = \frac{\partial l_k(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*)}{\partial \hat{p}_{1+k}},
S_{\partial l/\partial \hat{p}_{1+k}^*}(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*) = \frac{\partial l_k(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*)}{\partial \hat{p}_{1+k}^*}.
\]

To find the maximum likelihood estimates (MLEs) of \( D \), \( \hat{p}_{1+k} \), and \( \hat{p}_{1+k}^* \) under \( H_0 \) respectively, we need to solve the following 2\( K \) + 1 equations:

\[
\begin{align*}
S_{\partial l/\partial D}(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*) = 0, \\
S_{\partial l/\partial \hat{p}_{1+k}}(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*) = 0, \quad k = 1, 2, ..., K, \\
S_{\partial l/\partial \hat{p}_{1+k}^*}(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*) = 0, \quad k = 1, 2, ..., K.
\end{align*}
\]

Variances and covariances for the efficient scores are given by

\[
\begin{align*}
I_{\partial l/\partial D} & = \text{Var}(S_{\partial l/\partial D}(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*)), \\
I_{\partial l/\partial \hat{p}_{1+k}} & = \text{Var}(S_{\partial l/\partial \hat{p}_{1+k}}(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*)), \\
I_{\partial l/\partial \hat{p}_{1+k}^*} & = \text{Var}(S_{\partial l/\partial \hat{p}_{1+k}^*}(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*)).
\end{align*}
\]

Denote

\[
I_{\partial l/\partial D}^{-1} = I_{\partial l/\partial \hat{p}_{1+k}}^{-1} = I_{\partial l/\partial \hat{p}_{1+k}^*}^{-1}.
\]

Hence, the likelihood score test for the homogeneity hypothesis \( H_0 : D = \hat{D} \) is given by

\[
X^2 = \sum_{k=1}^{K} S_{\partial l/\partial D}^2(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*) = \left[ \sum_{k=1}^{K} S_{\partial l/\partial \hat{p}_{1+k}}^2(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*) \right] \left[ \sum_{k=1}^{K} S_{\partial l/\partial \hat{p}_{1+k}^*}^2(D, \hat{p}_{1+k}, \hat{p}_{1+k}^*) \right]^{-1}.
\]

Unfortunately, \( \hat{D} \), \( \hat{p}_{1+k} \), and \( \hat{p}_{1+k}^* \) cannot be expressed in a closed form and this makes the likelihood score test \( X^2 \) less appealing in practice. To overcome this issue, applying the theory of homogeneity score test extended to nuisance parameters [12] we propose the following modified score statistic

\[
X^{*2} = \sum_{k=1}^{K} S_{\partial l/\partial D}^2(D^*, \hat{p}_{1+k}^*, \hat{p}_{1+k}^{*+}) = \left[ \sum_{k=1}^{K} S_{\partial l/\partial \hat{p}_{1+k}}^2(D^*, \hat{p}_{1+k}^*, \hat{p}_{1+k}^{*+}) \right] \left[ \sum_{k=1}^{K} S_{\partial l/\partial \hat{p}_{1+k}^*}^2(D^*, \hat{p}_{1+k}^*, \hat{p}_{1+k}^{*+}) \right]^{-1},
\]

where \( D^* \), \( \hat{p}_{1+k}^* \) and \( \hat{p}_{1+k}^{*+} \) are any consistent estimators of \( D \), \( \hat{p}_{1+k} \), and \( \hat{p}_{1+k}^* \), respectively. To this end, we choose \( D^* \) to be

\[
\begin{align*}
D^* & = (D^* , D^* , D^* , ..., D^* )' \quad \text{and} \quad \hat{p}_{1+k}^* \quad \text{and} \quad \hat{p}_{1+k}^{*+} \quad \text{be the solutions to the following equations:}
\end{align*}
\]

or equivalently the following quartic polynomial equations,

\[
\begin{align*}
0 = a_0 + a_1 \rho_{1+k} + a_2 \rho_{1+k}^2 + a_3 \rho_{1+k}^3 + a_4 \rho_{1+k}^4, \\
0 = b_0 + b_1 \rho_{1+k} + b_2 \rho_{1+k}^2 + b_3 \rho_{1+k}^3 + b_4 \rho_{1+k}^4,
\end{align*}
\]

where
an asymptotic chi-square distribution with

and

efficiency for
general, it is not an efficient estimator to
the calculation of \( \Delta \) in (2) is quite tedious. None-

Here, \( D^* \) is analogous to the well-known Mantel-Haen-
szel estimator [13]. It is a consistent estimator to \( D \). In
general, it is not an efficient estimator to \( D \). The proof of
consistency and the conditions for achieving asymptotic efficiency for \( D^* \) is presented in Appendix. We notice that
the calculation of \( IkD_{p1+k}p_{1+k} \) in (2) is quite tedious. None-
thless, it is easy to show that \( IkD_{p1+k}p_{1+k} \) is simply given by

\[
n_k w_k = p_{1+k} \;
\]

with \( w_k(D, p_{1+k}) = p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} - 4D^2 \)
(see Appendix for the proof). It can be shown that \( X^2 \) has
an asymptotic chi-square distribution with \( K - 1 \) degrees of
freedom under \( H_0 \). Therefore, the homogeneity hypothe-
sis \( H_0 \) is rejected at level \( \alpha \) when \( X^2 \geq \chi^2_{K-1,1-\alpha} \), where
\( \chi^2_{K-1,1-\alpha} \) is the 100 \times (1 - \alpha) percentile point of the chi-
square distribution with \( K - 1 \) degrees of freedom. Finally,
it is noteworthy that if the consistent estimators of \( D_k, p_{1+k} \)
and \( p_{1+k} \) are the constrained MLEs under \( H_0 \) then the second
term of (2) vanishes, since

\[
\sum_{k=1}^{K} S_k D_{p1+k} p_{1+k} = 0,
\]
and (2) reduces to the likeli-

Asymptotic power and sample size
We will present the asymptotic power and sample size for-
formulae based on \( X^2 \) [14]. For this purpose, we assume \( n_k = \)
\( n_k^* \) for some \( n \) and \( n_k^* > 0 \). Let \( D_k, p_{1+k} \) and \( p_{1+k} \)
be the true parameter values for \( D_k, p_{1+k} \) and \( p_{1+k} \) under \( H_1 \), where
\( k = 1, 2, \ldots, K \) and \( D_k \neq D_j \) for at least a pair \( k \neq j \). Thus, the
asymptotic power for the homogeneity score test \( X^2 \) at \( \alpha \)
level is given by

\[
Pr(X^2 \geq \chi^2_{K-1,1-\alpha} | H_1) = Pr(\chi^2_{K-1,1} \geq \chi^2_{K-1,1-\alpha}),
\]

where \( \chi^2_{K-1,1} \) denotes the non-central chi-square distri-
bution with \( K - 1 \) degrees of freedom with the non-centrality
parameter being

\[
\Delta = n \sum_{k=1}^{K} \left[ n_k (p_{1+k} - D_k)^2 + (p_{1+k} - n_k D_{p1+k})^2 + \left( \frac{1}{w_k(D_{p1+k})} \right) \sum_{l=1}^{K} \left( \frac{p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k}}{p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k}} \right)^2 \right] - \frac{1}{w_k(D_{p1+k})} \sum_{l=1}^{K} \left( \frac{p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k}}{p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k}} \right),
\]

where

\[
d = \sum_{k=1}^{K} \left( \frac{p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k}}{p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k}} \right) \left( \frac{1}{w_k(D_{p1+k})} \sum_{l=1}^{K} \left( \frac{p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k}}{p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k} + p_{1+k} D_{p1+k}} \right)^2 \right).
\]

The desirable sample size \( n \) required to attain the power at
1 \(-\beta \) with \( D_k, p_{1+k} \) and \( p_{1+k} \) being the true parameter values
for \( D_k, p_{1+k} \) and \( p_{1+k} \) under the alternative \( H_1 \) at nomi-
inal level \( \alpha \) can be found by the relation

\[
\chi^2_{K-1,1,\beta} (\Delta) = \chi^2_{K-1,1-\alpha} (\Delta),
\]

where \( \chi^2_{K-1,1,\beta} (\Delta) \) is the 100 \times \beta percentile point of the non-central chi-square distribution with \( K - 1 \) degrees of freedom and non-centrality parameter \( \Delta \). The sample size \( n \) can be readily obtained by solving the above equation.
Availability and requirements
We have implemented the test procedures for computing our score statistic $X^2$* in a Matlab project. Project name: gametic disequilibrium homogeneity score test (GDHST); Project home page: http://math.nenu.edu.cn/jhguo/program.htm Operating system: Windows XP; Programming language: Matlab 6.1; Licence: GNU GPL.

Results
Simulation results
To evaluate the performance of our proposed homogeneity score test, we include the homogeneity test recommended by Weir [9] in our comparison study. The corresponding test statistic for homogeneity is given by

$$T^2 = \sum_{k=1}^{K} (n_k - 3)(z_k - \bar{z})^2,$$

where $K$ is the total number of strata, $n_k$ is the total gamete number in stratum $k$, $z_k = \frac{1}{2} \ln(\frac{1+n_k}{1-k})$ is the Fisher’s $z$ transformation with $n_k = \frac{nK}{k} \{k+1-k^2 + k+1-k\}$ and $(x_{01k}, x_{10k}, x_{11k})$ being the number of the gamete array in the $k$-th stratum, and $\bar{z}$ is the average of the $z_k$ values.

We investigate the performance of $X^2$* and $T^2$ in terms of type I error rate and power. For type I error rates, we consider both equal and unequal allele probabilities varying from 0.1 to 0.5 across ($K = 3$ and 5) strata with equal sample sizes ($n_k = 50$, 100 and 200) for $k = 1,\ldots,K$ and common disequilibrium ($D = \frac{1}{2} D_{\text{min},0}$ and $\frac{1}{2} D_{\text{max},0}$), where $D_{\text{min}} = \max\{D_{1,\text{min},1},\ldots,D_{K,\text{min},K}\}$, $D_{\text{max}} = \min\{D_{1,\text{max},1},\ldots,D_{K,\text{max},K}\}$.

Monte Carlo simulations with 5,000 repetitions at 0.05 nominal level are summarized in Table 1, 2, 3, 4. Table 1 shows the performance of empirical type I error rates for $X^2$* and $T^2$ with equal allele probabilities across $K = 3$ strata. We observe the following.

1. When $D$ is large (i.e., $\frac{1}{2} D_{\text{max}}$), both tests generally appear to be quite liberal (e.g., empirical size being 10 times of the nominal level), especially for small sample size (e.g., $nk = 50$) and small allele probability (e.g., $p_{11} = p_{11} = (0.1, 0.1, 0.1)$). Such liberty in empirical size is more severe in $T^2$ than in our asymptotic homogeneity test $X^2$* and is significantly alleviated in $X^2$* when sample size increases. However, sample size increase does not alleviate the liberty of $T^2$ much. In fact, even for $nk = 3200$ for $k = 1,2,3$, $T^2$ is still very liberal for $D = 0.045$ with empirical type I errors rate being 0.456 (data are not shown).

2. For other settings, both tests perform quite satisfactorily in the sense that their empirical sizes are well controlled around the pre-chosen nominal level. In general, the larger the sample size, the closer the empirical type I error rate to the pre-chosen nominal level.

Table 2 reports the empirical size performance of $X^2$* and $T^2$ for unequal allele probabilities across $K = 3$ strata. We observe similar phenomena above. However, our asymptotic homogeneity test $X^2$* performs quite well in all settings under consideration for moderate to large sample sizes (i.e., $n_k = 100$ and 200) while it is not the case for $T^2$. For $T^2$, the resultant empirical type I error rate can be extremely inflated even for large sample design (e.g., more than 17 times of the nominal level when $n_k = 200$ for $(k = 1,2,3)$).

For $T^2$, the resultant empirical type I error rate can be extremely inflated even for large sample design (e.g., more than 17 times of the nominal level when $n_k = 200$ for $(k = 1,2,3)$).

Table 3 and 4 shows the empirical type I error rate performance of $X^2$* and $T^2$ for $K = 5$. The parameter settings are similar to Table 1 and 2. According to the simulation results, liberty issue becomes more serious and larger sample sizes are required to attain similar performance when $K$ increases from 3 to 5 under similar parameter settings.

Since many type I error rates for $X^2$* and $T^2$ are liberal in Tables 1 to 4. The two-sided $t$-test is conducted to determined if an empirical type I error rate is significantly different from the nominal level of 0.05. The $t$-test statistics is

$$\sqrt{m-1} \frac{W-0.05}{\sqrt{W(1-W)}},$$

where $m = 5000$ and $W$ represents the empirical type I error rate of $X^2$* or $T^2$. Here, the $t$-test is almost identical to the $z$-test for the sample size is very large. Those empirical type I error rates which are significantly different from the nominal level of 0.05 are underlined in Tables 1 to 4. In Table 1, the total number of significant difference from the nominal level of 0.05 for $X^2$* and $T^2$ is 28 and 38, respectively. The pair $(28, 38)$ can be further decomposed to $(14, 14)$, $(8, 13)$ and $(6, 11)$ according to $n = 50$, 100 and 200. The decreasing rate of the number of empirical type I error rates which is significant different from the nominal level of 0.05 for $X^2$* is $14/18-6/18 = 44.4\%$ as $n$ increases from 50 to 200. While the corresponding decreasing rate for $T^2$ is $14/18-11/18 = 16.7\%$. It is easy to see that our $X^2$* is less liberal than $T^2$ as sample size increases.

Table 2 reports the empirical size performance of $X^2$* and $T^2$ for unequal allele probabilities across $K = 3$ strata. We observe similar phenomena above. However, our asymptotic homogeneity test $X^2$* performs quite well in all settings under consideration for moderate to large sample sizes (i.e., $n_k = 100$ and 200) while it is not the case for $T^2$. For $T^2$, the resultant empirical type I error rate can be extremely inflated even for large sample design (e.g., more than 17 times of the nominal level when $n_k = 200$ for $(k = 1,2,3)$).

Table 3 and 4 shows the empirical type I error rate performance of $X^2$* and $T^2$ for $K = 5$. The parameter settings are similar to Table 1 and 2. According to the simulation results, liberty issue becomes more serious and larger sample sizes are required to attain similar performance when $K$ increases from 3 to 5 under similar parameter settings.

Since many type I error rates for $X^2$* and $T^2$ are liberal in Tables 1 to 4. The two-sided $t$-test is conducted to determined if an empirical type I error rate is significantly different from the nominal level of 0.05. The $t$-test statistics is

$$\sqrt{m-1} \frac{W-0.05}{\sqrt{W(1-W)}},$$

where $m = 5000$ and $W$ represents the empirical type I error rate of $X^2$* or $T^2$. Here, the $t$-test is almost identical to the $z$-test for the sample size is very large. Those empirical type I error rates which are significantly different from the nominal level of 0.05 are underlined in Tables 1 to 4. In Table 1, the total number of significant difference from the nominal level of 0.05 for $X^2$* and $T^2$ is 28 and 38, respectively. The pair $(28, 38)$ can be further decomposed to $(14, 14)$, $(8, 13)$ and $(6, 11)$ according to $n = 50$, 100 and 200. The decreasing rate of the number of empirical type I error rates which is significant different from the nominal level of 0.05 for $X^2$* is $14/18-6/18 = 44.4\%$ as $n$ increases from 50 to 200. While the corresponding decreasing rate for $T^2$ is $14/18-11/18 = 16.7\%$. It is easy to see that our $X^2$* is less liberal than $T^2$ as sample size increases.
### Table 1: Empirical type I error rates for $X^2$ and $T^2$ for equal allele probabilities across $K=3$ strata under $H_0$

| n      | $D$    | $p_{1-}$ | $p_{11}$ | $X^2$ | $T^2$ |
|--------|--------|----------|----------|-------|-------|
| 50, 50, 50 | -0.125 | 0.5, 0.5, 0.5 | 0.5, 0.5, 0.5 | 0.047 | 0.110 |
|        | 0.0    |          |          | 0.053 | 0.052 |
|        | 0.125  |          |          | 0.076 | 0.116 |
|        | -0.075 | 0.5, 0.5, 0.5 | 0.3, 0.3, 0.3 | 0.051 | 0.061 |
|        | 0.0    |          |          | 0.053 | 0.049 |
|        | 0.075  |          |          | 0.075 | 0.063 |
|        | -0.045 | 0.3, 0.3, 0.3 | 0.3, 0.3, 0.3 | 0.043 | 0.021 |
|        | 0.0    |          |          | 0.044 | 0.053 |
|        | 0.105  |          |          | 0.100 | 0.167 |
|        | -0.025 | 0.5, 0.5, 0.5 | 0.1, 0.1, 0.1 | 0.061 | 0.036 |
|        | 0.0    |          |          | 0.047 | 0.041 |
|        | 0.025  |          |          | 0.089 | 0.029 |
|        | -0.015 | 0.3, 0.3, 0.3 | 0.1, 0.1, 0.1 | 0.024 | 0.017 |
|        | 0.0    |          |          | 0.025 | 0.047 |
|        | 0.035  |          |          | 0.104 | 0.117 |
|        | -0.005 | 0.1, 0.1, 0.1 | 0.1, 0.1, 0.1 | 0.031 | 0.024 |
|        | 0.0    |          |          | 0.024 | 0.087 |
|        | 0.045  |          |          | 0.413 | 0.515 |
| 100, 100, 100 | -0.125 | 0.5, 0.5, 0.5 | 0.5, 0.5, 0.5 | 0.049 | 0.106 |
|        | 0.0    |          |          | 0.052 | 0.051 |
|        | 0.125  |          |          | 0.065 | 0.112 |
|        | -0.075 | 0.5, 0.5, 0.5 | 0.3, 0.3, 0.3 | 0.051 | 0.059 |
|        | 0.0    |          |          | 0.049 | 0.051 |
|        | 0.075  |          |          | 0.055 | 0.063 |
|        | -0.045 | 0.3, 0.3, 0.3 | 0.3, 0.3, 0.3 | 0.046 | 0.024 |
|        | 0.0    |          |          | 0.048 | 0.051 |
|        | 0.105  |          |          | 0.048 | 0.156 |
|        | -0.025 | 0.5, 0.5, 0.5 | 0.1, 0.1, 0.1 | 0.053 | 0.029 |
|        | 0.0    |          |          | 0.048 | 0.047 |
|        | 0.025  |          |          | 0.089 | 0.030 |
|        | -0.015 | 0.3, 0.3, 0.3 | 0.1, 0.1, 0.1 | 0.026 | 0.013 |
|        | 0.0    |          |          | 0.029 | 0.048 |
|        | 0.035  |          |          | 0.075 | 0.109 |
|        | -0.005 | 0.1, 0.1, 0.1 | 0.1, 0.1, 0.1 | 0.012 | 0.014 |
|        | 0.0    |          |          | 0.013 | 0.057 |
|        | 0.045  |          |          | 0.278 | 0.474 |
| 200, 200, 200 | -0.125 | 0.5, 0.5, 0.5 | 0.5, 0.5, 0.5 | 0.050 | 0.050 |
|        | 0.0    |          |          | 0.052 | 0.048 |
|        | 0.125  |          |          | 0.058 | 0.112 |
|        | -0.075 | 0.5, 0.5, 0.5 | 0.3, 0.3, 0.3 | 0.050 | 0.057 |
|        | 0.0    |          |          | 0.050 | 0.049 |
|        | 0.075  |          |          | 0.054 | 0.058 |
|        | -0.045 | 0.3, 0.3, 0.3 | 0.3, 0.3, 0.3 | 0.049 | 0.025 |
|        | 0.0    |          |          | 0.049 | 0.049 |
|        | 0.105  |          |          | 0.052 | 0.156 |
|        | -0.025 | 0.5, 0.5, 0.5 | 0.1, 0.1, 0.1 | 0.053 | 0.030 |
|        | 0.0    |          |          | 0.050 | 0.051 |
|        | 0.025  |          |          | 0.052 | 0.032 |
|        | -0.015 | 0.3, 0.3, 0.3 | 0.1, 0.1, 0.1 | 0.044 | 0.014 |
|        | 0.0    |          |          | 0.044 | 0.051 |
|        | 0.035  |          |          | 0.045 | 0.105 |
|        | -0.005 | 0.1, 0.1, 0.1 | 0.1, 0.1, 0.1 | 0.020 | 0.010 |
|        | 0.0    |          |          | 0.020 | 0.049 |
|        | 0.045  |          |          | 0.098 | 0.463 |

The empirical type I error rate which is significant for the two-sided $t$-test at the nominal level of 0.05 is marked with underline.
Table 2: Empirical type I error rates for $X^2$ and $T^2$ for unequal allele probabilities across $K = 3$ strata under $H_0$

| $n$           | $D$  | $p_{1.}$  | $p_{2.}$  | $X^2$  | $T^2$  |
|--------------|------|-----------|-----------|--------|--------|
| 50, 50, 50   | -0.045 | 0.5, 0.4, 0.3 | 0.5, 0.4, 0.3 | 0.048  | 0.044  |
|              | 0.0   | 0.049     | 0.051     |        |        |
|              | 0.105 | 0.071     | 0.149     |        |        |
|              | -0.015 | 0.5, 0.4, 0.3 | 0.5, 0.3, 0.1 | 0.046  | 0.037  |
|              | 0.0   | 0.046     | 0.051     |        |        |
|              | 0.035 | 0.066     | 0.105     |        |        |
|              | -0.005 | 0.5, 0.3, 0.1 | 0.5, 0.3, 0.1 | 0.063  | 0.036  |
|              | 0.0   | 0.053     | 0.052     |        |        |
|              | 0.045 | 0.100     | 0.474     |        |        |
|              | -0.015 | 0.5, 0.4, 0.3 | 0.3, 0.2, 0.1 | 0.040  | 0.032  |
|              | 0.0   | 0.042     | 0.049     |        |        |
|              | 0.035 | 0.074     | 0.101     |        |        |
|              | -0.005 | 0.5, 0.3, 0.1 | 0.3, 0.2, 0.1 | 0.056  | 0.033  |
|              | 0.0   | 0.048     | 0.054     |        |        |
|              | 0.045 | 0.108     | 0.452     |        |        |
|              | -0.005 | 0.3, 0.2, 0.1 | 0.3, 0.2, 0.1 | 0.048  | 0.031  |
|              | 0.0   | 0.041     | 0.053     |        |        |
|              | 0.045 | 0.123     | 0.452     |        |        |
| 100, 100, 100| -0.045 | 0.5, 0.4, 0.3 | 0.5, 0.4, 0.3 | 0.051  | 0.048  |
|              | 0.0   | 0.050     | 0.050     |        |        |
|              | 0.105 | 0.055     | 0.162     |        |        |
|              | -0.015 | 0.5, 0.4, 0.3 | 0.5, 0.3, 0.1 | 0.047  | 0.043  |
|              | 0.0   | 0.046     | 0.050     |        |        |
|              | 0.035 | 0.054     | 0.150     |        |        |
|              | -0.005 | 0.5, 0.3, 0.1 | 0.5, 0.3, 0.1 | 0.064  | 0.037  |
|              | 0.0   | 0.052     | 0.051     |        |        |
|              | 0.045 | 0.059     | 0.658     |        |        |
|              | -0.015 | 0.5, 0.4, 0.3 | 0.3, 0.2, 0.1 | 0.044  | 0.035  |
|              | 0.0   | 0.047     | 0.051     |        |        |
|              | 0.035 | 0.051     | 0.124     |        |        |
|              | -0.005 | 0.5, 0.3, 0.1 | 0.3, 0.2, 0.1 | 0.055  | 0.033  |
|              | 0.0   | 0.052     | 0.053     |        |        |
|              | 0.045 | 0.063     | 0.623     |        |        |
|              | -0.005 | 0.3, 0.2, 0.1 | 0.3, 0.2, 0.1 | 0.055  | 0.033  |
|              | 0.0   | 0.043     | 0.051     |        |        |
|              | 0.045 | 0.061     | 0.593     |        |        |
| 200, 200, 200| -0.045 | 0.5, 0.4, 0.3 | 0.5, 0.4, 0.3 | 0.050  | 0.052  |
|              | 0.0   | 0.050     | 0.052     |        |        |
|              | 0.105 | 0.053     | 0.211     |        |        |
|              | -0.015 | 0.5, 0.4, 0.3 | 0.5, 0.3, 0.1 | 0.049  | 0.049  |
|              | 0.0   | 0.048     | 0.049     |        |        |
|              | 0.035 | 0.049     | 0.220     |        |        |
|              | -0.005 | 0.5, 0.3, 0.1 | 0.5, 0.3, 0.1 | 0.058  | 0.037  |
|              | 0.0   | 0.051     | 0.050     |        |        |
|              | 0.045 | 0.048     | 0.860     |        |        |
|              | -0.015 | 0.5, 0.4, 0.3 | 0.3, 0.2, 0.1 | 0.048  | 0.043  |
|              | 0.0   | 0.049     | 0.048     |        |        |
|              | 0.035 | 0.050     | 0.190     |        |        |
|              | -0.005 | 0.5, 0.3, 0.1 | 0.3, 0.2, 0.1 | 0.056  | 0.035  |
|              | 0.0   | 0.051     | 0.051     |        |        |
|              | 0.045 | 0.053     | 0.829     |        |        |
|              | -0.005 | 0.3, 0.2, 0.1 | 0.3, 0.2, 0.1 | 0.051  | 0.034  |
|              | 0.0   | 0.050     | 0.050     |        |        |
|              | 0.045 | 0.049     | 0.791     |        |        |

The empirical type I error rate which is significant for the two-sided $t$-test at the nominal level of 0.05 is marked with underline.
Table 3: Empirical type I error rates for $X^2*$ and $T^2$ for equal allele probabilities across $K = 5$ strata under $H_0$

| n           | D     | $p_{1+}$       | $p_{+1}$       | $X^2*$ | $T^2$ |
|-------------|-------|----------------|----------------|--------|-------|
| 50, 50, 50, 50, 50 | -0.125 | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.041 | 0.146 |
| 0.0         |       | 0.053          | 0.048          |        |       |
| 0.125       |       | 0.090          | 0.148          |        |       |
| -0.075      | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.049 | 0.059 |
| 0.0         |       | 0.053          | 0.051          |        |       |
| 0.075       | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.036 | 0.018 |
| -0.045      | 0.0   | 0.039          | 0.040          |        |       |
| 0.0         |       | 0.172          | 0.238          |        |       |
| 0.105       | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.075 | 0.029 |
| 0.0         |       | 0.071          | 0.038          |        |       |
| 0.025       | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.015 | 0.012 |
| 0.0         |       | 0.028          | 0.043          |        |       |
| 0.035       | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.025 | 0.024 |
| -0.005      | 0.0   | 0.026          | 0.079          |        |       |
| 0.0         |       | 0.500          | 0.715          |        |       |
| 0.045       | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.046 | 0.135 |
| 100, 100, 100, 100, 100 | -0.125 | 0.051          | 0.050          |        |       |
| 0.0         |       | 0.068          | 0.141          |        |       |
| 0.125       | 0.051 | 0.051          | 0.051          |        |       |
| -0.075      | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.050 | 0.059 |
| 0.0         |       | 0.051          | 0.051          |        |       |
| 0.075       | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.046 | 0.020 |
| -0.045      | 0.0   | 0.044          | 0.052          |        |       |
| 0.0         |       | 0.051          | 0.212          |        |       |
| 0.105       | 0.0   | 0.058          | 0.048          |        |       |
| -0.025      | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.063 | 0.030 |
| 0.0         |       | 0.057          | 0.029          |        |       |
| 0.025       | 0.0   | 0.057          | 0.029          |        |       |
| -0.015      | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.026 | 0.009 |
| 0.0         |       | 0.026          | 0.009          |        |       |
| 0.035       | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.016 | 0.007 |
| -0.005      | 0.0   | 0.008          | 0.006          |        |       |
| 0.0         |       | 0.476          | 0.667          |        |       |
| 200, 200, 200, 200, 200 | -0.125 | 0.049          | 0.049          |        |       |
| 0.0         |       | 0.049          | 0.049          |        |       |
| 0.125       | 0.061 | 0.133          |                |        |       |
| -0.075      | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.051 | 0.055 |
| 0.0         |       | 0.051          | 0.051          |        |       |
| 0.075       | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.3, 0.3, 0.3, 0.3, 0.3 | 0.053 | 0.022 |
| -0.045      | 0.0   | 0.050          | 0.051          |        |       |
| 0.0         |       | 0.050          | 0.203          |        |       |
| 0.105       | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.054 | 0.027 |
| 0.0         |       | 0.048          | 0.049          |        |       |
| 0.025       | 0.0   | 0.053          | 0.028          |        |       |
| -0.015      | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.037 | 0.008 |
| 0.0         |       | 0.037          | 0.049          |        |       |
| 0.035       | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.1, 0.1, 0.1, 0.1, 0.1 | 0.017 | 0.007 |
| -0.005      | 0.0   | 0.018          | 0.053          |        |       |
| 0.0         |       | 0.018          | 0.651          |        |       |
| 0.045       | 0.193 | 0.651          |                |        |       |

The empirical type I error rate which is significant for the two-sided t-test at the nominal level of 0.05 is marked with underline.
Table 4: Empirical type I error rates for $\chi^2$ and $T^2$ for unequal allele probabilities across $K = 5$ strata under $H_0$

| n           | $D$     | $p_{1+}$ | $p_{2+}$ | $\chi^2$ | $T^2$ |
|-------------|---------|----------|----------|----------|-------|
| 50, 50, 50, 50 | -0.045  | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.048   | 0.049 |
|             | 0.0     | 0.085    | 0.163    | 0.047    | 0.041 |
|             | -0.025  | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.057    | 0.054 |
|             | 0.0     | 0.077    | 0.060    | 0.042    | 0.057 |
|             | 0.005   | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.048    | 0.035 |
|             | 0.0     | 0.133    | 0.489    | 0.045    | 0.037 |
|             | 0.0045  | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.045    | 0.053 |
|             | 0.0     | 0.082    | 0.100    | 0.032    | 0.023 |
|             | -0.015  | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.034    | 0.048 |
|             | 0.0     | 0.104    | 0.136    | 0.045    | 0.034 |
|             | 0.035   | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.035    | 0.057 |
|             | 0.0     | 0.160    | 0.475    | 0.110    | 0.230 |
|             | 0.0045  | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.050    | 0.051 |
|             | 0.0     | 0.060    | 0.190    | 0.052    | 0.053 |
|             | -0.025  | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.050    | 0.050 |
|             | 0.0     | 0.051    | 0.051    | 0.063    | 0.067 |
|             | 0.005   | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.050    | 0.037 |
|             | 0.0     | 0.041    | 0.049    | 0.058    | 0.053 |
|             | 0.0045  | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.046    | 0.040 |
|             | 0.0     | 0.047    | 0.051    | 0.054    | 0.050 |
|             | -0.015  | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.037    | 0.023 |
|             | 0.0     | 0.040    | 0.168    | 0.047    | 0.035 |
|             | 0.035   | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.038    | 0.053 |
|             | 0.0     | 0.060    | 0.610    | 0.060    | 0.610 |
|             | 0.0045  | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.049    | 0.050 |
|             | 0.0     | 0.048    | 0.049    | 0.053    | 0.230 |
|             | -0.025  | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.049    | 0.062 |
|             | 0.0     | 0.051    | 0.050    | 0.053    | 0.081 |
|             | 0.025   | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.051    | 0.039 |
|             | 0.0     | 0.043    | 0.048    | 0.052    | 0.085 |
|             | 0.0045  | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.048    | 0.044 |
|             | 0.0     | 0.049    | 0.049    | 0.048    | 0.169 |
|             | -0.015  | 0.5, 0.4, 0.3, 0.2, 0.1 | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.044    | 0.029 |
|             | 0.0     | 0.045    | 0.053    | 0.045    | 0.229 |
|             | 0.035   | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.3, 0.2, 0.1, 0.2, 0.3 | 0.047    | 0.035 |
|             | 0.0     | 0.040    | 0.053    | 0.051    | 0.815 |

The empirical type I error rate which is significant for the two-sided t-test at the nominal level of 0.05 is marked with underline.
For Table 2, the total number of significant difference from the nominal level of 0.05 for $X^2*$ and $T^2$ is 17 and 33, respectively. The pair (17, 33) can again be decomposed to (14, 14), (8, 13) and (6, 11) according to $n = 50, 100$ and 200. The decreasing rate of the number of empirical type I error rates which is significant different from the nominal level of 0.05 for $X^2*$ is $10/18 - 1/18 = 50.0\%$ as $n$ increases from 50 to 200. While the decreasing rate for $T^2$ is $12/18 - 10/18 = 11.1\%$. The decreasing rate of our $X^2*$ is again more significant than that of $T^2$.

In Table 3 to 4, the strata increases from 3 to 5. However, the decreasing rates of the number of empirical type I error rates which is significant different from the nominal level of 0.05 for Tables 3 and 4 is very close to that of Tables 1 and 2, respectively. Therefore, we have reason to believe that this decreasing rate is not greatly affected by the number of strata.

Table 5 summarizes the empirical powers for $X^2*$ and $T^2$. Here, $\{D_k\}$ are specified under $H_1$ and we set $D_k = D_0 + \delta (k - 1)$. For $K = 3$, we consider: (i) $D_0 = -0.03, \delta = 0.03$ and (ii) $D_0 = -0.05, \delta = 0.05$. For $K = 5$, we consider: (i) $D_0 = -0.06, \delta = 0.03$ and (ii) $D_0 = -0.1, \delta = 0.05$. From the simulation results, we observe both $X^2*$ and $T^2$ perform similarly under the designed parameter settings. In general, powers increase with $n$ and $\delta$.

In view of the above results, we prefer the proposed homogeneity test $X^2*$ to the traditional $T^2$ which is based on the Fisher’s test of homogeneity among correlation coefficient.

### Real and hypothetical examples

It is reported that mutations at the cystic fibrosis transmembrane conductance regulator gene (CFTR) cause cystic fibrosis, the most prevalent severe genetic disorder.

| n          | D               | $p_{1+}$ | $p_{1-}$ | $X^2*$ | $T^2$ |
|------------|-----------------|----------|----------|--------|-------|
| 50, 50, 50 | -0.03, 0.0, 0.03| 0.5, 0.5, 0.5 | 0.5, 0.5, 0.5 | 0.205 | 0.210 |
| 100, 100, 100 | 0.311 | 0.306 |
| 200, 200, 200 | 0.560 | 0.558 |
| 50, 50, 50 | -0.03, 0.0, 0.03 | 0.5, 0.4, 0.3 | 0.5, 0.4, 0.3 | 0.196 | 0.203 |
| 100, 100, 100 | 0.360 | 0.368 |
| 200, 200, 200 | 0.630 | 0.641 |
| 50, 50, 50 | -0.03, 0.0, 0.03 | 0.5, 0.5, 0.5 | 0.5, 0.4, 0.3 | 0.197 | 0.187 |
| 100, 100, 100 | 0.354 | 0.340 |
| 200, 200, 200 | 0.612 | 0.612 |
| 50, 50, 50 | -0.05, 0.0, 0.05 | 0.5, 0.5, 0.5 | 0.5, 0.5, 0.5 | 0.430 | 0.421 |
| 100, 100, 100 | 0.724 | 0.720 |
| 200, 200, 200 | 0.958 | 0.958 |
| 50, 50, 50 | -0.05, 0.0, 0.05 | 0.5, 0.4, 0.3 | 0.5, 0.4, 0.3 | 0.474 | 0.483 |
| 100, 100, 100 | 0.797 | 0.806 |
| 200, 200, 200 | 0.980 | 0.981 |
| 50, 50, 50 | -0.05, 0.0, 0.05 | 0.5, 0.5, 0.5 | 0.5, 0.4, 0.3 | 0.457 | 0.446 |
| 100, 100, 100 | 0.763 | 0.760 |
| 200, 200, 200 | 0.973 | 0.973 |
| 50, 50, 50, 50, 50 | -0.06, -0.03, 0.0, 0.0, 0.03, 0.06 | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.523 | 0.512 |
| 100, 100, 100, 100, 100 | 0.846 | 0.841 |
| 200, 200, 200, 200, 200 | 0.993 | 0.993 |
| 50, 50, 50, 50, 50 | -0.06, -0.03, 0.0, 0.0, 0.03, 0.06 | 0.5, 0.4, 0.3, 0.4, 0.5, 0.5 | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.526 | 0.526 |
| 100, 100, 100, 100, 100 | 0.854 | 0.853 |
| 200, 200, 200, 200, 200 | 0.995 | 0.995 |
| 50, 50, 50, 50, 50 | -0.06, -0.03, 0.0, 0.0, 0.03, 0.06 | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.535 | 0.522 |
| 100, 100, 100, 100, 100 | 0.855 | 0.850 |
| 200, 200, 200, 200, 200 | 0.994 | 0.994 |
| 50, 50, 50, 50, 50 | -0.1, -0.05, 0.0, 0.0, 0.05, 0.1 | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.957 | 0.953 |
| 100, 100, 100, 100, 100 | 1.000 | 1.000 |
| 200, 200, 200, 200, 200 | 1.000 | 1.000 |
| 50, 50, 50, 50, 50 | -0.1, -0.05, 0.0, 0.0, 0.05, 0.1 | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.960 | 0.960 |
| 100, 100, 100, 100, 100 | 1.000 | 1.000 |
| 200, 200, 200, 200, 200 | 1.000 | 1.000 |
| 50, 50, 50, 50, 50 | -0.1, -0.05, 0.0, 0.0, 0.05, 0.1 | 0.5, 0.5, 0.5, 0.5, 0.5 | 0.5, 0.4, 0.3, 0.4, 0.5 | 0.957 | 0.954 |
| 100, 100, 100, 100, 100 | 1.000 | 1.000 |
| 200, 200, 200, 200, 200 | 1.000 | 1.000 |
in individuals of European descent. Mateu [15] conducted a worldwide genetic analysis of the CFTR region and analyzed normal allele and haplotype variation at two single-nucleotide polymorphisms (SNPs), namely the T854/AvaII (2694 T/G) and TUB20/PVUII (4006-200 G/A). The T854 and TUB20 markers can be used to define the core haplotypes since they are diallelic, have presumably much lower mutation rates than the other polymorphisms and the ancestral state can be inferred for them.

Mateu [15] reported the T854-TUB20 haplotype frequencies by 18 populations. After communicating with one of their coauthors (Prof. Kenneth, pers. comm. 1996), it was found that their reported gametic frequencies were actually the maximum likelihood estimates of the gametic probabilities obtained from HAPLO, a software which can be applicable to missing data. In other words, all individuals with results for at least one of the two markers were included to estimate the gametic frequencies and no actual gametic counts were available. To create the gametic counts for each population, we first estimate the total number of participants in each population by the number of individuals who yielded results for at least one of the two markers. The reported gametic frequencies of each population given in Mateu [15] are multiplied to the estimated number of participants of this population and the closest integers are then taken to be the estimated gametic counts. The estimated gametic counts across the 18 populations are reported in Table 6, which is adopted as the real data in all subsequent analysis.

It is noticed that the gametic counts for the populations of Japanese (14th) and Surui (18th) are (0, 32, 0, 12)' and (0, 7, 0, 35)', respectively and their estimated gametic disequilibrium \( D_{k}, D_{k, \min} \) and \( D_{k, \max} \) are all equal to zero. Therefore, we will exclude these two populations for subsequent homogeneity testings. We consider the following scenarios.

(i) Homogeneity of gametic disequilibrium among the 16 populations (i.e., excluding Japanese and Surui). The statistic value of our proposed \( X^2* \) is 121.35 with \( p \)-value being less than 0.0001 while that of \( T^2 \) yields 99.64 with \( p \)-value being less than 0.0001. In this case, both tests reject the homogeneity hypothesis at the 0.05 nominal level.

(ii) Homogeneity of gametic disequilibrium among those populations with the same numbers of participants for both markers T854 and TUB20 (i.e., Mbuti, Yemenites, Druze, Adygei, Catalans, Basques, Chinese, and Nasioi).

Table 6: T854-TUB20 haplotype counts by 18 populations and some related statistics

| Population | Gametic counts | Allele frequencies | Disequilibrium | Estimation |
|------------|----------------|--------------------|----------------|------------|
|            | 1 - 1 | 1 - 2 | 2 - 1 | 2 - 2 | 1 | 1 | r | D' | D |
| Africa:    |        |        |        |        |        |        | | | |
| Biaka      | 5 | 16 | 12 | 29 | 0.339 | 0.274 | -0.058 | -0.132 | -0.012 |
| Mbuti      | 0 | 14 | 5 | 14 | 0.424 | 0.152 | -0.363 | -1.000 | -0.064 |
| Tanzanian  | 0 | 13 | 3 | 20 | 0.361 | 0.083 | -0.227 | -1.000 | -0.030 |
| North Africa: |        |        |        |        |        |        | | | |
| Saharawi   | 5 | 22 | 12 | 16 | 0.491 | 0.309 | -0.263 | -0.401 | -0.061 |
| Middle East: |        |        |        |        |        |        | | | |
| Yemenites  | 2 | 29 | 4 | 5 | 0.775 | 0.150 | -0.444 | -0.570 | -0.066 |
| Druze      | 2 | 47 | 10 | 4 | 0.778 | 0.191 | -0.713 | -0.786 | -0.116 |
| Europe:    |        |        |        |        |        |        | | | |
| Adygei     | 1 | 34 | 9 | 5 | 0.714 | 0.204 | -0.689 | -0.860 | -0.125 |
| Russians   | 0 | 17 | 10 | 5 | 0.531 | 0.313 | -0.718 | -1.000 | -0.166 |
| Finns      | 0 | 23 | 6 | 4 | 0.697 | 0.182 | -0.715 | -1.000 | -0.127 |
| Catalans   | 3 | 53 | 18 | 9 | 0.675 | 0.253 | -0.661 | -0.788 | -0.135 |
| Basques    | 4 | 72 | 15 | 17 | 0.704 | 0.176 | -0.500 | -0.701 | -0.087 |
| Asia:      |        |        |        |        |        |        | | | |
| Kazakhs    | 1 | 18 | 2 | 12 | 0.576 | 0.091 | -0.155 | -0.421 | -0.022 |
| Chinese    | 0 | 22 | 1 | 20 | 0.512 | 0.023 | -0.158 | -1.000 | -0.012 |
| Japanese   | 0 | 32 | 0 | 12 | 0.727 | 0 | NaN | NaN | 0 |
| Yakut      | 0 | 18 | 1 | 4 | 0.783 | 0.044 | -0.405 | -1.000 | -0.034 |
| Pacific:   |        |        |        |        |        |        | | | |
| Nasioi     | 1 | 20 | 0 | 22 | 0.488 | 0.023 | 0.158 | 1.000 | 0.012 |
| America:   |        |        |        |        |        |        | | | |
| Maya       | 2 | 15 | 0 | 31 | 0.354 | 0.042 | 0.282 | 1.000 | 0.027 |
| Surui      | 0 | 7 | 0 | 35 | 0.167 | 0 | NaN | NaN | 0 |
Our proposed statistic $X^2^*$ yields 50.56 with $p$-value being less than 0.0001 while $T^2$ gives 39.72 with $p$-value being less than 0.0001. Again, both tests suggest rejection of the homogeneity hypothesis at the 0.05 nominal level. Suppose that another research team wants to re-conduct the same genetic analysis. In this regard, it is sensible to ask, "How large is the sample size for each population in order to achieve, say, 90% power at the 0.05 nominal level?"

(iii) Homogeneity of gametic disequilibrium among those populations in Europe.

Our statistic $X^2^*$ yields 7.48 with $p$-value being 0.11 and $T^2$ yields 7.26 with $p$-value being 0.12. Both tests do not reject the homogeneity hypothesis at the 0.05 nominal level. In this case, we have evidence to believe that populations in Europe reach their gametic equilibrium.

To end this section, we analyze the hypothetical example of gametic disequilibrium between two loci (A, B) in ten populations described in Zapata and Alvarez [8]. Here, the gametic counts are simply set by multiplying the haplotype frequencies given in Zapata and Alvarez [8] by 1000. The data are reproduced in Table 7. Obviously, the $r$ values are homogeneous across the ten populations. For $D^*$ values, Zapata and Alvarez [8] utilized the bias-corrected non-parametric bootstrap method to obtain the 95% confidence interval for each $D^*$ value. Observing that the resultant confidence intervals have no intersection, they concluded that $D^*$ are heterogeneous. They suggested tests for homogeneity of gametic disequilibrium should be based on $D^*$, whose range is allele probability independent, rather than $r$. Although, the $D$ values in Table 7 seem to be homogeneous, our homogeneity score test yields $X^2^* = 33.44$ with $p$-value being less than 0.0001. Therefore, our test procedure also suggests the rejection of the homogeneity of gametic disequilibrium across the ten populations. In this case, our test reaches the same conclusion drawn by Zapata and Alvarez [8].

### Discussion

Verification of the homogeneity assumption of gametic disequilibrium across several populations is crucial in gametic disequilibrium analysis. We note that traditional homogeneity test on gametic disequilibrium is based on the Fisher’s test of homogeneity among correlation coefficients. However, our simulations demonstrate that this traditional test may not perform satisfactorily. Specifically, it can be very conservative or liberal, for almost all the cases in which the common true gametic disequilibrium $D$ is bounded away from zero. Most importantly, these kinds of conservativeness and liberty can not effectively alleviated with increased sample sizes.

Our proposed large-sample homogeneity score test on gametic disequilibrium across several independent populations requires the count of haplotypes as input. In practice, only genotype data can be obtained in most situations. To employ our method, one can use some haplotyping software, such as PHASE, HAPLOTYPER, to resolve the genotype data as haplotype data. In this way, it separates haplotype phasing and gametic disequilibrium homogeneity test. Naturally, it is more promising to extend our method which can directly handle the genotype data. In this sense, model assumptions are based on genotype data. However, the haplotype phase uncertainty for the double heterozygotes makes the definition of

| Population | $A_0 B_0$ | $A_0 B_1$ | $A_1 B_0$ | $A_1 B_1$ | $A_0$ | $B_0$ | $D^*$ | $D$ |
|------------|-----------|-----------|-----------|-----------|-------|-------|------|-----|
| 1          | 495       | 405       | 5         | 95        | 0.90  | 0.50  | 0.300| 0.900| 0.045|
| 2          | 540       | 360       | 1         | 90        | 0.90  | 0.55  | 0.300| 0.814| 0.045|
| 3          | 479       | 371       | 21        | 129       | 0.85  | 0.50  | 0.300| 0.714| 0.054|
| 4          | 460       | 340       | 40        | 160       | 0.80  | 0.50  | 0.300| 0.600| 0.060|
| 5          | 671       | 229       | 29        | 71        | 0.90  | 0.70  | 0.300| 0.589| 0.041|
| 6          | 539       | 261       | 61        | 139       | 0.80  | 0.60  | 0.300| 0.490| 0.059|
| 7          | 615       | 185       | 85        | 115       | 0.80  | 0.70  | 0.300| 0.393| 0.055|
| 8          | 373       | 227       | 127       | 273       | 0.60  | 0.50  | 0.300| 0.367| 0.073|
| 9          | 403       | 197       | 147       | 253       | 0.60  | 0.55  | 0.300| 0.332| 0.073|
| 10         | 325       | 175       | 175       | 325       | 0.50  | 0.50  | 0.300| 0.300| 0.075|

Note: The table above represents a hypothetical example of gametic disequilibrium between two loci (A, B) with two alleles ($A_0$, $A_1$ and $B_0$, $B_1$, respectively) across ten populations.
consistent estimate of $D$. Let $w_k = w_k(D, p_{1+k}, p_{1+k})$, $v_k = 1/(p_{01+k} + p_{10+k})$. Thus, the asymptotic variance of $D^*$ under $H_0$ is given by

$$\text{AsyVar}(D^*) = \frac{(\sum_{k=1}^{K} w_k v_k^2 / b_k)}{n(\sum_{k=1}^{K} v_k^2)^2}.$$  

Let the information matrix with respect to $D$, $p_{1+k}$ and $p_{1+k}$ under $H_0$ be

$$I = \begin{pmatrix}
I_{1Dp_{1+i}} & I_{1Dp_{1+i}} & \cdots & I_{1Dp_{1+iK}} \\
I_{1Dp_{1+i}} & I_{1Dp_{1+i}} & \cdots & I_{1Dp_{1+iK}} \\
\vdots & \vdots & \cdots & \vdots \\
I_{KDp_{1+iK}} & \cdots & \cdots & I_{KDp_{1+iK}}
\end{pmatrix}$$

By inverting the information matrix $I$, we can obtain the asymptotic variance of $\hat{D}$, that is,

$$\text{AsyVar}(D) = \frac{1}{n} \left( \sum_{k=1}^{K} \frac{b_k}{w_k} \right)^{-1}.$$  

By Cauchy-Schwarz inequality

$$\left( \sum_{k=1}^{K} v_k \right)^2 \leq \left( \sum_{k=1}^{K} w_k v_k^2 / b_k \right) \sum_{k=1}^{K} w_k v_k^2 / b_k,$$

we have

$$\text{AsyVar}(\hat{D}) = \text{AsyVar}(D^*).$$  

To this end, we obtain the sufficient and necessary condition for the asymptotic efficiency of $D^*$, that is, $w_k v_k = c, k = 1, 2, \ldots, K$, where $c$ is a constant independent of all parameters. When $D = 0$, the condition is satisfied. From this, we know that $D^*$ is inefficient for general cases.

**A simple expression for** $I_{kDp_{1+i}p_{1+i}}$

For the $k$-th stratum, denote the information matrix with respect to $D_{y_k} p_{1+k}$ and $p_{1+k}$ by

$$I_k = \begin{pmatrix}
I_{kD_{y_k}D_{y_k}} & I_{kD_{y_k}p_{1+k}} & I_{kD_{y_k}p_{1+k}} \\
I_{kD_{p_{1+k}p_{1+k}}} & I_{kD_{p_{1+k}p_{1+k}}} & I_{kD_{p_{1+k}p_{1+k}}} \\
I_{kD_{p_{1+k}}p_{1+k}} & I_{kD_{p_{1+k}}p_{1+k}} & I_{kD_{p_{1+k}}p_{1+k}}
\end{pmatrix}$$

According to the property of inverse matrix, $I_{kD_{y_k}p_{1+i}p_{1+i}} (D_{y_k} p_{1+k} p_{1+k})$ is equal to the reciprocal of the $(1, 1)$ element of $I_k^{-1}$. By the property of MLEs, we have

$$\sqrt{n_k}(D - D_{y_k} p_{1+k} - p_{1+k} p_{1+k} - p_{1+k})^d \rightarrow N(0, n_k I_k^{-1}(D_{y_k} p_{1+k} p_{1+k})).$$
where $\hat{D}_k, \hat{p}_{1+k} = x_{1+k}/n_k$ and $\hat{p}_{+1k} = x_{+1k}/n_k$ are the MLEs of $D_k, p_{1+k}$ and $p_{+1k}$ respectively. Hence, the asymptotic variance of $\sqrt{n_k}{\hat{D}_k}$ is $n_k/I_{k\hat{D}_k}\hat{p}_{1+k}\hat{p}_{+1k}(D_k, p_{1+k}, p_{+1k})$. On the contrary, by the Central Limit Theorem, $\sqrt{n_k}(y_k - g_k)$ follows an asymptotic normal distribution $N(0, \Sigma_k)$. By $\delta$ method, we immediately get that $\sqrt{n_k}(D_k - \hat{D}_k)$ follows an asymptotic normal distribution $N(0, w_k(D_k, p_{1+k}, p_{+1k}))$. Therefore, we can obtain the exact expression $I_{k\hat{D}_k}\hat{p}_{1+k}\hat{p}_{+1k}(D_k, p_{1+k}, p_{+1k}) = n_k/w_k(D_k, p_{1+k}, p_{+1k})$.

Naturally, the expression of $I_{k\hat{D}_k}\hat{p}_{1+k}\hat{p}_{+1k}(D, p_{1+k}, p_{+1k})$ is just $I_{k\hat{D}}\hat{p}_{1+k}\hat{p}_{+1k}(D, p_{1+k}, p_{+1k})$ by substituting $D$ for $D_k$.

Acknowledgements
This research was supported by the National Natural Science Foundation of China (Grant Numbers 10431010 and 10701022), National 973 Key Project of China (2007CB11002), NCET-04-0310, EYTP, the Jilin Distinguished Young Scholars Program (Grant Number 20030113) and the Program Innovative Research Team (PCSIRT) in University (#IR070519). The work of ML Tang was fully supported by a grant from the Research Grant Council of the Hong Kong Special Administration (Project no. HKBU261007).

References
1. Lewontin RC: The genetic basis of evolutionary change New York: Columbia University Press; 1974.
2. Jorde LB: Linkage disequilibrium as a gene mapping tool. Am J Hum Genet 1995, 56:1-14.
3. Hedrick PW, Jain S, Holden L: Multilocus systems in evolution. Evol Biol 1978, 11:1-182.
4. Weir BS: Inferences about linkage disequilibrium. Biometrics 1979, 35:235-254.
5. Hedrick PW: Gametic disequilibrium measures: proceed with caution. Genetics 1987, 117:331-341.
6. Mueller JC: Linkage disequilibrium for different scales and applications. Brief Bioinform 2004, 5:355-364.
7. Lewontin RC, Kojima K: The evolutionary dynamics of complex polymorphisms. Evolution 1960, 14:458-472.
8. Zapata C, Alvarez G: Testing for homogeneity of gametic disequilibrium among populations. Evolution 1997, 51:606-607.
9. Weir BS: Genetic Data Analysis II Sunderland, Massachusetts: Sinauer Associates; 1996.
10. Fisher RA: Statistical methods for research workers New York: Oliver and Boyd; 1925.
11. Lewontin RC: The interaction of selection and linkage. I. General considerations; heterotic models. Genetics 1964, 49:49-67.
12. Tarone RE: Homogeneity score tests with nuisance parameters. Commun Stat-Theor M 1988, 17:1549-1556.
13. Mantel N, Haenszel W: Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959, 22(4):719-748.
14. Guo JH, Ma YP, Shi NZ, Lau TS: Testing for homogeneity of relative difference under inverse sampling. Comput Stat Data Anal 2004, 44:613-624.
15. Mateu E, Calafell F, Lao O, Batsheva BT, Kidd JR, Pakstis A, Kidd KK, Bertranpetit J: Worldwide genetic analysis of the CFTR region. Am J Hum Genet 2001, 68:103-117.