Published in final edited form as:
Genet Epidemiol. 2011 September ; 35(6): 515–525. doi:10.1002/gepi.20601.

Detection of *cis*-acting regulatory SNPs using allelic expression data

Rui Xiao[1,2] and **Laura J. Scott**[2]

[1]Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine

[2]Department of Biostatistics and Center for Statistical Genetics, University of Michigan

Abstract

Allelic expression (AE) imbalance between the two alleles of a gene can be used to detect *cis*-acting regulatory SNPs (rSNPs) in individuals heterozygous for a transcribed SNP (tSNP). In this paper, we propose three tests for AE analysis focusing on phase-unknown data and any degree of linkage disequilibrium (LD) between the rSNP and tSNP: a test based on the minimum p-value of a one-sided F and two-sided t tests proposed previously for phase-unknown data, a test that combines these two p-values, and a mixture-model based test. We compare these three tests to the F and t tests and an existing regression-based test for phase-known data. We show that the ranking of the tests based on power depends most strongly on the magnitude of the LD between the rSNP and tSNP. For phase-unknown data we find that under a range of scenarios, our proposed tests have higher power than the F and t tests when LD between the rSNP and tSNP is moderate (~.2 < D'_{RT} < ~.8). We further demonstrate that the presence of a second ungenotyped rSNP almost never invalidates the proposed tests nor substantially changes their power rankings. For detection of *cis*-acting regulatory SNPs using phase-unknown AE data, we recommend the F test when the rSNP and tSNP are in or near linkage equilibrium (D'_{RT} < .2); the t test when the two SNPs are in strong LD (D'_{RT} > .7); and the mixture-model based test for intermediate LD levels (.2 < D'_{RT} < .7).

Keywords

allelic expression imbalance; *cis*-acting regulatory SNP; linkage phase; cDNA; genomic DNA; mixture model; likelihood ratio test; parametric bootstrap

Introduction

mRNA levels are affected by environmental variation, epigenetic modifications, and genetic regulatory elements that reside within and outside of the mRNA transcript [Gilad et al., 2008; Cheung and Spielman, 2009; Pastinen 2010]. *Trans*-acting regulatory elements regulate both alleles of the gene equally and can be located on the same or a different chromosome [Monks et al., 2004; Cheung et al., 2005]. *Cis*-acting regulatory elements regulate the expression of the gene on the same chromosome and are often, but not always,

Address for correspondence: Rui Xiao, Ph.D., Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, 19104-6021, Phone: (215) 746-4474, FAX: (215) 573-1050, rxiao@mail.med.upenn.edu.

in close proximity to the gene they regulate [Yan et al., 2002; Bray et al., 2004; Monks et al., 2004; Cheung et al., 2005; Stranger et al., 2005]. Global identification of *cis*-acting regulatory variants (rSNP) enables an understanding of variants that influence local gene expression [Ge et al., 2009; Pickerel et al., 2010; Cheung et al., 2010] and can aid with identification of causative SNPs and genes in regions identified by genome-wide association studies [Cookson et al., 2009; Speliotes et al., 2010; Fogarty et al., 2010].

One way to detect *cis*-acting variants is via analysis of mRNA from primary or transformed cells. Following reverse transcription of RNA to cDNA, the relative levels of the allele specific transcript in the cDNA can be measured by genotyping of a transcribed SNP (tSNP) or by RNA-seq. *Cis*-acting regulatory variants cause unequal levels of the transcripts from the two alleles of the gene (allelic expression (AE) imbalance), which can be detected by comparing the levels of the two transcribed alleles in individuals heterozygous for the tSNP and often quantitated as the allelic expression ratio (AER). Each transcribed allele serves as an internal standard for the other to control for *trans*-regulatory and environmental factors that affect the expression of both alleles.

A variety of approaches exist to test for AE-rSNP association although each is limited in scope of application. Many AE studies compare the AER in cDNA to the AER in genomic DNA (gDNA) from the same samples, with the gDNA as a reference for equal levels of the two transcript alleles [Bray et al., 2004; Campino et al., 2008; Pant et al., 2006; Fogarty et al., 2010]. When a single rSNP is in r^2=1 with the tSNP, the underlying AER will be, on average, consistently higher or lower than the gDNA level (Figure 1A) and a t test comparing the mean AER between cDNA and gDNA or comparing the mean AER of cDNA normalized by gDNA to 1 can be performed [Bray et al., 2004; Campino et al., 2008]. The use of gDNA as a reference assumes that any technical bias in measurement of AER for the cDNA is the same for the gDNA.

When the rSNP is not in $r^2 = 1$ with the tSNP, the distribution of the AER will depend on the rSNP-tSNP haplotypes present in the study sample. For samples with known rSNP-tSNP haplotypes (rSNP-tSNP phase-known data), a regression-based test for AE-rSNP association can be used [Campino et al., 2008; Ge et al., 2009], in which the haplogenotype of the rSNP homozygotes is coded as intermediate to the two haplogenotypes of the double heterozygotes (for example, $\frac{\mathrm{rT}}{\mathrm{Rt}}=-1, \frac{\mathrm{RT}}{\mathrm{Rt}}=\frac{\mathrm{rT}}{\mathrm{rt}}=0, \frac{\mathrm{RT}}{\mathrm{rt}}=1$).

However, in practice, rSNP-tSNP phase is often unknown in a given set of samples. The existing tests for phase-unknown data are designed to work optimally when D'_{RT} is low or relatively high [Fogarty et al., 2010]. When D'_{RT} =1 and r^2 <1, there is only one rSNP-tSNP haplogenotype configuration present in the rSNP heterozygotes (Figure 1B) and the mean AER of the rSNP heterozygotes can be compared either to the mean AER in gDNA (as described above) or to the mean AER of the rSNP homozyogtes using a two-sided t test [Bray and O'Donovan, 2004; Fogarty et al., 2010]. The cDNA of rSNP homozygotes can also serve as a reference for equal levels of the two tSNP alleles. Because the AER is measured from the cDNA rather than the gDNA, the potential bias of AER in the reference group may be reduced compared to using the gDNA as the reference, although fewer reference samples may be available. When $D'_{RT} < 1$, there are two possible haplogenotypes for rSNP heterozygotes and, relative to the rSNP homozygotes, we expect to observe one cluster of samples with higher AER and another cluster with lower AER (Figure 1C). We previously proposed a one-sided F test for higher AE variance in rSNP heterozygotes than in rSNP homozygotes. The power of the test is maximal when the rSNP and tSNP are in linkage equilibrium (LE) [Fogarty et al., 2010]. Teare et al. [2006] proposed a four-component mixture model and expectation-maximization (EM) algorithm to analyze AE data and a likelihood ratio test (LRT) to compare mean AE in rSNP heterozygotes and homozygotes. However, assessment of the significance of the LRT is not described and the usual chi-square distribution cannot be used due to non-identifiability of parameters in the finite mixture model [Hartigan, 1985]. The lack of tests designed for intermediate LD range ($0 < D'_{RT} < 1$) has limited the set of rSNP and tSNP combinations that can be effectively tested.

Our goal in this paper is to develop tests to detect *cis*-acting regulatory SNPs in the context of genotype or RT-PCR data when $D'_{RT} < 1$ and linkage phase is unknown, although the ideas we describe can be extended to RNA-seq data. We describe three statistical tests that seek to capture better the available information from the AE data distribution: a test based on the minimum p-value of the F and t tests, a test that combines these two p-values, and a mixture-model based test which fits a two-component mixture model for rSNP heterozygotes. For the minimum- and combined-p-value tests, we use permutation to assess significance allowing for non-normality or correlated tests, while for the mixture-model based test we employ the parametric bootstrap. We evaluate the performance of the three new tests relative to the existing F and t tests and a regression-based test for phase-known data.

We demonstrate through computer simulation that the F test is generally the most powerful test when the two SNPs are in LE or low LD ($D'_{RT} < .2$), but has fairly low power when the two SNPs are in high LD ($D'_{RT} > .5$). In contrast, the t test generally is the least powerful test when LD is low ($D'_{RT} < .2$), but most powerful when LD is high ($D'_{RT} > .5$). When LD is intermediate, the mixture-model based test generally has the highest power, slightly higher than the combined-p-value test. We also demonstrate that the presence of a second ungenotyped rSNP generally does not invalidate these tests, but may result in reduced or increased power, depending on the LD structure between the three loci and the direction of effect of the two rSNPs.

Methods

Model and assumptions

We initially assume that the differential expression of a gene is caused in part by a single *cis*-acting rSNP with alleles R and r, with R causing higher expression of the allele on its chromosome compared to r. AE imbalance is measured in N independent individuals who are heterozygous for a tSNP with alleles T and t. Let p_R and p_T denote the frequencies of R and T, and D'_{RT} the standardized LD between the two SNPs. For individual i, let $G_i \in \{$RR, Rr, rr} be the genotype of the rSNP and $H_i \in \left\{\frac{\mathrm{RT}}{\mathrm{rt}}, \frac{\mathrm{rT}}{\mathrm{Rt}}, \frac{\mathrm{RT}}{\mathrm{Rt}}, \frac{\mathrm{rT}}{\mathrm{rt}}\right\}$ be the haplogenotype of the rSNP and tSNP.

We define the AER as the ratio of the allele T transcript level to the allele t transcript level, and use the natural logarithm of this AER normalized by the corresponding ratio in gDNA for the tSNP heterozygotes as the outcome variable

$$y = \ln \mathrm{AER}_{\mathrm{cDNA}} - \mathrm{mean}(\ln {}^{\mathrm{T}}\!/_{\mathrm{t}})_{\mathrm{gDNA\_\ Tt}} \quad (1)$$

In what follows, we will refer to *y* as lnAER. Normalization of lnAER by the gDNA mean (ln T/t) does not affect the type I error rate or power of the tests we propose, but may control for any systematic differences in quantification of the two tSNP alleles, and thus allows for interpretation of the estimated AE imbalance effect size.

Compared to rSNP homozygotes ($h = \frac{\mathrm{RT}}{\mathrm{Rt}}, \frac{\mathrm{rT}}{\mathrm{rt}}$) for which we do not expect to observe AE imbalance, in the presence of AE imbalance, Rr heterozygotes will show an increased T:t expression ratio if $h = \frac{\mathrm{RT}}{\mathrm{rt}}$ and a decreased T:t expression ratio if $h = \frac{\mathrm{rT}}{\mathrm{Rt}}$.

For individual i with haplogenotype *h*, we assume y_i is normally distributed with mean μ_h and variance σ^2, where

$$\mu_h = \begin{cases} \mu_0 & \text{for} h = \frac{\mathrm{RT}}{\mathrm{Rt}} \text{or} \frac{\mathrm{rT}}{\mathrm{rt}} \\ \mu_0 + \alpha_{\mathrm{R}} & \text{for} h = \frac{\mathrm{RT}}{\mathrm{rt}} \\ \mu_0 - \alpha_{\mathrm{R}} & \text{for} h = \frac{\mathrm{rT}}{\mathrm{Rt}} \end{cases} \quad (2)$$

Under the null hypothesis of no AE imbalance, $\alpha_R = 0$. We assume that there is no difference in the mean or variance of *y* between the RR and rr homozygotes.

Minimum- and combined-p-value tests based on existing F and t tests

When the rSNP and tSNP are in LE or low LD ($D'_{RT} < .2$), the two RrTt haplogenotypes ($h = \frac{\mathrm{RT}}{\mathrm{rt}}, \frac{\mathrm{rT}}{\mathrm{Rt}}$) have similar frequencies. In the presence of AE imbalance, we expect approximately half the Rr heterozygotes to have high lnAER and the remainder to have low lnAER, resulting in an increased lnAER variance for Rr heterozygotes compared to the combined RR and rr homozygotes. For this situation, we [Fogarty et al., 2010] proposed using the F test for equal variances against the one-sided alternative as a test for AE imbalance.

When the rSNP and tSNP are in moderate to high LD ($D'_{RT} > .3$), one of the two RrTt haplogenotypes ($h = \frac{\mathrm{RT}}{\mathrm{rt}}, \frac{\mathrm{rT}}{\mathrm{Rt}}$) is substantially more common than the other. In the presence of AE imbalance, we expect mean lnAER for Rr heterozygotes to be higher or lower than for the combined RR and rr homozygotes, depending on which haplogenotype is more common. We [Fogarty et al., 2010] proposed using a (two-sided) two-sample t test for the hypothesis that mean lnAER of the Rr heterozygotes differs from that of the combined RR and rr homozygotes, allowing for unequal variances between the heterozygous and homozygous groups due to the mixing distribution for the Rr heterozygotes.

For both these tests, we used permutations of the rSNP genotypes to assess significance while accounting for violation of the normality assumption due to the nature of the mixed distribution of the lnAER in the rSNP heterozygotes.

The F test tends to be more powerful when the rSNP and tSNP are in low LD ($D'_{RT} < .2$) and the t test tends to be more powerful given high LD ($D'_{RT} > .7$) (see Results). To take advantage of the strengths of both of the two tests, we consider two additional tests. The minimum-p-value test

$$\mathrm{T_{min}}=\min(\mathrm{P_F},\mathrm{P_t}) \quad (3)$$

selects the minimum of the p-values for the F and t tests (P_t, P_F) while the combined-p-value test

$$\mathrm{T_{com}}=-2(\ln\mathrm{P_F}+\ln\mathrm{P_t}) \quad (4)$$

uses Fisher's (1948) method to meta-analyze the information from the two tests. We again use permutation of the rSNP genotypes to assess significance for the minimum- and combined-p-value tests to account for the dependence of the F and t tests.

Mixture-model based test

Given unknown linkage phase and incomplete LD, the lnAER data follow a mixture distribution. We therefore propose a mixture-model based test which fits a two-component normal mixture model for the rSNP heterozygotes, with likelihood:

$$\mathrm{L}=\prod_{\mathrm{i}\in\{G_\mathrm{i}=\mathrm{RR,rr}\}} f(y_\mathrm{i};\mu_0,\sigma^2)\times\prod_{\mathrm{i}\in\{G_\mathrm{i}=\mathrm{Rr}\}}\left\{\pi f(y_\mathrm{i};\mu_0+\alpha_\mathrm{R},\sigma^2)+(1-\pi)f(y_\mathrm{i};\mu_0-\alpha_\mathrm{R},\sigma^2)\right\} \quad (5)$$

Here, $f(\mu, \sigma^2)$ is the density function for a normal distribution with mean μ and variance σ^2 and π is the mixing proportion.

We perform a likelihood ratio test (LRT) of the null hypothesis based on the likelihood ratio statistic

$$\Lambda=-2\ln\frac{L(\hat{\theta}_0)}{L(\hat{\theta}_1)} \quad (6)$$

where $\hat{\theta}_0 = (\hat{\mu}_0, \hat{\sigma}^2)$ and $\hat{\theta}_1 = (\tilde{\pi}, \tilde{\mu}_0, \tilde{\alpha}_R, \tilde{\sigma}^2)$ are the maximum likelihood estimators (MLEs) under the null and alternative hypotheses, respectively. Since the likelihood cannot be maximized analytically, we obtain MLEs by the simplex method [Nelder and Mead, 1965]. To assess significance for Λ, we apply the parametric bootstrap [McLachlan, 1987], since the chi-square distribution cannot be used to approximate the null distribution of LRT in finite mixture models [Hartigan, 1985]. For each bootstrap, we simulate the lnAER data from the distribution with parameters estimated under null hypothesis, and calculate the LRT statistic based on the bootstrapped data. We estimate the p-value as the proportion of the bootstrap LRT statistics greater than the observed LRT statistic; no ties were observed.

Phase known regression test

We compare power of the five tests for phase unknown data to that of an existing regression-based test [Ge et al., 2009] for phase-known data, in which the lnAER data are regressed on the haplogenotype coded as an additive model: 0 for one of the heterozygous haplogenotyes ($h = \frac{\mathrm{rT}}{\mathrm{Rt}}$), 1 for the combined homozygous haplogenotyes ($h = \frac{\mathrm{RT}}{\mathrm{Rt}}, \frac{\mathrm{rT}}{\mathrm{rt}}$), and 2 for the other heterozygous haplogenotype ($h = \frac{\mathrm{RT}}{\mathrm{rt}}$).

Simulations: one regulatory SNP

We evaluated the performance of the tests to detect association between AE imbalance and the potential rSNP by simulating samples with varying numbers of Tt heterozygotes N, allele frequencies p_{R} and p_{T}, $\mathrm{D'}_{\mathrm{RT}}$ values, and mean lnAER effect α_{R} with fixed variance $\sigma^2 = 1$. For each individual, we simulated haplotype pairs according to the conditional probabilities of the two-locus haplogenotypes assuming ascertainment for Tt heterozygotes. For example,

$$f_{h=\frac{\underline{\mathrm{RT}}}{\underline{\mathrm{Rt}}}} = \frac{p(\underline{\mathrm{RT}}, \underline{\mathrm{Rt}})}{p(\mathrm{Tt})} = \frac{w_{\underline{\mathrm{RT}}} w_{\underline{\mathrm{Rt}}}}{p_{\mathrm{T}}(1 - p_{\mathrm{T}})} \quad (7)$$

where w_l is the frequency of haplotype $l \in \{\underline{\mathrm{RT}}, \underline{\mathrm{rT}}, \underline{\mathrm{Rt}}, \underline{\mathrm{rt}}\}$. We then simulated the corresponding lnAER data from a normal distribution with the appropriate haplogenotype-specific mean described in (2). We chose the value of α_{R} for a given N to yield informative power comparisons between the tests.

Simulations: two regulatory SNPs

So far, we have assumed a single rSNP. In fact, there could be more than one [see for example Ge et al., 2009]. To assess the impact of a second (ungenotyped) regulatory SNP on the power and relative rankings of the proposed tests, we simulated lnAER data assuming there is a second *cis*-acting rSNP with alleles $\mathrm{R_U}$ and $\mathrm{r_U}$ influencing allelic expression, where $p_{\mathrm{R_U}}$ is the frequency of the allele causing higher expression.

Given two regulatory SNPs $\mathrm{R_G}$ (genotyped) and $\mathrm{R_U}$ (ungenotyped), there are 16 possible haplogenotypes for Tt heterozygotes. Probabilities for these haplogenotypes can be calculated as a function of the pairwise D' values $\mathrm{D'}_{\mathrm{R_G R_U}}$, $\mathrm{D'}_{\mathrm{R_G T}}$ and $\mathrm{D'}_{\mathrm{R_U T}}$, and the third-order LD $\mathrm{D}_{\mathrm{R_G R_U T}}$ between the three loci [Bennett 1954]:

$$\mathrm{D}_{\mathrm{R_G R_U T}} = w_{\mathrm{R_G R_U T}} - p_{\mathrm{T}} \mathrm{D}_{\mathrm{R_G R_U}} - p_{\mathrm{R_U}} \mathrm{D}_{\mathrm{R_G T}} - p_{\mathrm{R_G}} \mathrm{D}_{\mathrm{R_U T}} - p_{\mathrm{R_G}} p_{\mathrm{R_U}} p_{\mathrm{T}} \quad (8)$$

Here $w_{\mathrm{R_G R_U T}}$ is the haplotype frequency, and $\mathrm{D}_{\mathrm{R_G R_U}}$, $\mathrm{D}_{\mathrm{R_G T}}$ and $\mathrm{D}_{\mathrm{R_U T}}$ are the unnormalized pairwise LD for the three loci. The normalized third order LD

$$D'_{R_G R_U T} = \frac{D_{R_G R_U T} - D_{R_G R_U T}(\min)}{D_{R_G R_U T}(\max) - D_{R_G R_U T}(\min)} \quad (9)$$

[Thomson and Baur, 1984], where $D_{R_G R_U T}(\min)$ and $D_{R_G R_U T}(\max)$ are the lower and upper bounds for $D_{R_G R_U T}$.

We assume that the R_U allele of the ungenotyped rSNP increases mean lnAER by α_{R_U}, and that the two regulatory SNPs act additively, resulting in the pattern displayed by a "balloon plot" in Figure 2. In a balloon plot, the diameter of each circle corresponds to the frequency of the haplogenotype(s) to its right while the center of the circle corresponds to mean lnAER in individuals with that (those) halplogenotype(s). For example, lnAER for genotyped rSNP $R_G R_G$ homozygotes may display three clusters, with means $\mu_0 + \alpha_{R_U}$ (corresponding to haplogenotype $k=\frac{R_G R_U T}{R_G r_U t}$), $\mu_0 (k=\frac{R_G R_U T}{R_G R_U t}, \frac{R_G r_U T}{R_G r_U t})$, and $\mu_0 - \alpha_{R_U} (k=\frac{R_G r_U T}{R_G R_U t})$. As many as three clusters also may be present for $r_G r_G$ individuals, and six for $R_G r_G$ heterozygotes.

As in the one-rSNP case, for each individual, we simulate haplotype pairs based on probabilities analogous to those in (7), and the corresponding lnAER data with appropriate haplogenotype-specific mean.

Results

One rSNP

We first examined the type I error rates for the five tests that allow for the analysis of phase-unknown data: the F and t tests, the minimum- and combined-p-value tests, and the mixture-model based test for AE-rSNP association. We also included a regression-based test that requires phase-known data [Ge et al., 2009]. Our simulations show that type I error rate estimates are consistent with nominal significance levels α = .10, .05, and .01 (data not shown).

We evaluated the power of the six tests at significance level α = .05 as a function of LD levels between the regulatory and transcribed SNPs (D'_{RT}), and allele frequencies p_R and p_T. We first considered scenarios in which p_R is greater than or less than p_T (Figure 3A, 3C). We observed that the phase-known test has higher power for all settings investigated than the five phase-unknown tests and particularly when D'_{RT} is low, as expected. Figure 3A shows results for $p_R > p_T$, where $N = 100$, $\alpha_R = .85$, $p_R = .3$, and $p_T = .1$. Among the five tests for phase-unknown data, the F test has highest power when the tSNP and rSNP are in LE or low LD ($D'_{RT} < .2$), but the F test power decreases rapidly as D'_{RT} increases and is fairly low when D'_{RT} is moderate to high (> .4). When D'_{RT} is low, on average ~½ the rSNP heterozygotes have high lnAER and ~½ have low lnAER, resulting in a higher variance for the rSNP heterozygotes compared to the rSNP homozygotes (balloon plot of Figure 3A). When D'_{RT} is high, the variances are similar between the two rSNP genotype groups (balloon plot of Figure 3A). The t test is the least powerful test when D'_{RT} is low (< .4), but its power increases rapidly as D'_{RT} increases, and it becomes most powerful when D'_{RT} is high (> .7). When $D'_{RT} < .2$, the mean lnAER is similar in rSNP heterozygotes and homozygotes (balloon plot of Figure 3A). In contrast, for higher D'_{RT}, most rSNP heterozygotes will have either higher (when $h=\frac{\mathrm{RT}}{\mathrm{rt}}$ is the more common haplogenotype) or lower (when $h=\frac{\mathrm{rT}}{\mathrm{Rt}}$ is more common) lnAER compared to the rSNP homozygotes (balloon plot of Figure 3A).

The minimum- and combined-p-value tests are more powerful than both the F and t tests for moderate LD ($.3 < D'_{RT} < .5$), only slightly less powerful than the F test when LD is low ($D'_{RT} < .3$) or t test when LD is high ($D'_{RT} > .5$). The mixture-model based test shows similar performance as the minimum- and combined-p-value tests, but is the most powerful test among all five phase-unknown tests for moderate LD ($.3 < D'_{RT} < .7$) (Figure 3A). At moderate D'_{RT} the p-value based tests have higher power than the individual F or t test because they make use of information about the differences in both the AER mean and variance between the rSNP heterozygotes and homozygotes. The mixture-model based test explicitly acknowledges the mixed distribution of the AER in the two RrTt haplogenotypes and thus captures the information from mid range of D'_{RT} that the other tests for AE imbalance fail to do so.

Figure 3C shows results for $p_R < p_T$, where $p_R = .05$ and $p_T = .1$. When LD is moderate or high ($D'_{RT} \geq .3$), the shape of the power curves and the ranking of the tests based on power are similar to those observed in Figure 3A ($p_R > p_T$). However, when LD is low ($D'_{RT} < .3$), the F test is not more powerful than the mixture-model or p-value based tests, in contrast to the pattern observed in Figure 3A. When the rSNP is rare and the rSNP and tSNP are in low LD, the number of rSNP heterozygotes is very small (balloon plot of Figure 3C) and therefore the power for all tests decreases and particularly for the F test. Consequently, at every D'_{RT} level the mixture-model based test has the highest or nearly the highest power among the phase-unknown tests.

When p_R and p_T are similar or equal (Figure 3B), the rankings of the tests based on power are similar to those observed when $p_R \neq p_T$. However, the shapes of the power curves differ for all but the F test. The power of the four remaining phase-unknown tests and the phase-known test display an increasing-and-then-decreasing trend with power maximized at intermediate D'_{RT}, in contrast to a monotonically increasing trend when $p_R \neq p_T$. Power is reduced for high D'_{RT} because only a few tSNP heterozygotes are rSNP homozygotes (balloon plot of Figure 3B). The small number of rSNP homozygotes results in low power for the t test and consequently for the minimum- and combined-p-value tests, and also causes decreased power for the mixture-model based test due to poor estimation of μ_0.

We evaluated the impact of the number of tSNP heterozygotes N and rSNP AE imbalance effect size α_R on power by choosing combinations of $N = 50$, and 500 and α_R of 0.3 to 1.2 to allow for informative comparisons between the tests. We found that the power of the tests varies by scenario, but the rankings of the tests based on power remain largely consistent for different (N, α_R) combinations across different levels of LD (data not shown).

We compared the number of tSNP heterozygotes *N* necessary to obtain similar power levels for the most powerful phase-unknown test(s) at a given D'_{RT} to the phase-known test [Ge et al., 2009] by iteratively increasing *N* to achieve the desired power level (Table 1). We found that at moderately high D'_{RT}, similar or only slightly increased sample sizes are sufficient to achieve similar power in tests of phase-unknown and phase-known data. At lower D'_{RT}, substantially larger sample sizes are needed to achieve similar power.

Two rSNPs

We investigated the impact of a second (ungenotyped) rSNP on the type I error rate and power of the six tests to detect AE imbalance association with the genotyped putative rSNP (Figure 4). For type I error rate, we assumed that the genotyped putative rSNP has no effect on lnAER ($\alpha_{R_G} = 0$) while the ungenotyped rSNP has mean effect size $\alpha_{R_U} = .85$. For power, we assumed the two rSNPs have same effect size with mean $\alpha_{R_G} = \alpha_{R_U} = .85$ and act additively on gene expression, and we initially assumed that the minor alleles of the two rSNPs both increase gene expression. Figure 4 displays the type I error rates and power evaluated for different LD structures between the two rSNPs and the tSNP, assuming the allele frequencies for the genotyped and ungenotyped rSNPs $p_{R_G} = p_{R_U} = .3$, and tSNP $p_T = .1$.

Ungenotyped rSNP in LE with genotyped putative rSNP and tSNP

When the ungenotyped rSNP is in LE with both the genotyped putative rSNP and the tSNP ($D'_{R_G R_U} = D'_{R_U T} = 0$, $D'_{R_G T}$ varies from 0 to 1), empirical type I error rates are consistent with nominal expectation for $\alpha = .05$ (Figure 4A), .10 and .01 (data not shown); the ungenotyped rSNP has simply added noise but no bias (balloon plot of Figure 4A). For power, we found that when the ungenotyped rSNP is in LE with both the genotyped rSNP and the tSNP ($D'_{R_G R_U} = D'_{R_U T} = 0$), the rankings of the tests based on power are essentially unchanged compared to the single rSNP case, although the power of each test decreases slightly (compare Figures 3A and 4A). The presence of the second ungenotyped rSNP increases variation of the lnAER data for tSNP heterozygotes (balloon plot of Figure 4A).

Ungenotyped rSNP in LD with genotyped putative rSNP and tSNP

We next explored scenarios in which the ungentoyped rSNP is in moderate $D'_{R_G R_U} = D'_{R_U T} = .5$) to strong ($D'_{R_G R_U} = .5$, $D'_{R_U T} = 1$) LD with the genotyped rSNP and the tSNP. We oriented the two rSNPs such that the minor alleles of the two rSNPs are more likely to be on the same haplotype when the two rSNPs are in LD, and consequently the AE imbalance effects of the two rSNPs will add together. We observed both higher and lower type I error rates for the six tests than the nominal expectation across the range of $D'_{R_G T}$ (Figure 4B and 4C). For each test, the type I error rates are often higher than the nominal expectation when $D'_{R_G T}$ is closer to 0 or 1 because the difference in means or the variances between the $R_G r_G$ heterozygotes and the combined $R_G R_G$ and $r_G r_G$ homozygotes are higher due to the effect of the ungenotyped rSNP. The one-sided F test has a smaller than expected type I error rate when the genotyped putative rSNP is in moderate to high LD with the tSNP (Figure 4B, 4C); the ungenotyped rSNP causes the variance of the combined $R_G R_G$ and $r_G r_G$ homozygotes to be larger than that of the $R_G r_G$ heterozygote (balloon plot of Figure 4B, 4C). The relative rankings of the tests based on power are similar to those observed in the single rSNP scenario. However, the power of the tests is slightly higher than the single rSNP scenario, because of the LD between the ungenotyped rSNP with the tSNP and the consistent direction of AE imbalance effect of the two rSNPs. This power increase is more substantial when LD between the ungenotyped rSNP and the tSNP is stronger (Figure 4C).

If the two rSNPs act additively but the minor alleles of the two rSNPs regulate gene expression in opposite directions, power of all tests is slightly lower than for the single rSNP scenario when the two rSNPs are in low LD, and much lower when in moderate to high LD (data not shown).

Discussion

Cis-acting regulatory SNPs can be detected through measurement of the relative expression levels of the two alleles of a gene [Yan et al., 2002; Pastinen, 2010]. When $D'_{RT} < 1$, tests for AE imbalance can be carried out in phase-known data such as the HapMap CEU samples [Ge et al., 2009], or for phase-unknown samples [Fogarty et al., 2010], although few studies have chosen to evaluate these SNP pairs, likely owing to the lack of well-evaluated methods.

We have proposed three tests for AE-rSNP association that can be used for phase-unknown data, and compared their performance with our previously proposed F and t tests [Fogarty et al., 2010] designed for low and high D'_{RT} levels, respectively. The one-sided F test tends to be most powerful when the rSNP and tSNP are in LE or low D'_{RT}, and the two-sided t test when the two SNPs are in high LD. To take advantage of the differing strengths of the F and t tests, we propose the minimum- and combined-p-value tests. These tests tend to be more powerful than the F and t tests for moderate LD levels, and only slightly less powerful than the F test for low LD or the t test for high LD levels. Our mixture-model based test provides a single testing procedure alternative to the other four tests. We applied a two-component normal mixture model for the rSNP heterozygotes $\frac{\mathrm{RT}}{\mathrm{rt}}$ and $\frac{\mathrm{rT}}{\mathrm{Rt}}$ to model the mixed nature of the AE data. The performance of the mixture-model based test is similar to or slightly better than the minimum- and combined-p-value tests, although it requires the use of more complex model and analysis.

Although no one test has maximal power for all scenarios we have considered, in practice, we can determine the most likely powerful test(s) based on the sample size, allele frequencies of the rSNP and tSNP, estimated D' between them (either from the study sample or some other public data source such as HapMap samples), the variance of AER observed in the rSNP homozygotes, and the expected AE imbalance effect size [Fogarty et al., 2010].

Teare et al. [2006] proposed an alternative four-component mixture-model based method for AER-SNP, with components corresponding to the two rSNP heterozygous haplogenotypes $\frac{\mathrm{RT}}{\mathrm{rt}}$ and $\frac{\mathrm{rT}}{\mathrm{Rt}}$ and the two rSNP homozygous haplogenotypes $\frac{\mathrm{RT}}{\mathrm{Rt}}$ and $\frac{\mathrm{rT}}{\mathrm{rt}}$. They used a likelihood ratio test (LRT) comparing the four-component model to a one-component model given no AE imbalance, and compared the resulting LRT statistic to a chi-squared distribution on one degree of freedom [Mauro Santibánez Koref, personal communication]. This method has been used to estimate AER-SNP association in recent studies [Cunnington et al., 2010; Santibánez Koref et al., 2010]. However, the finite mixture model belongs to a non-regular parametric family and most classical asymptotic results do not apply, so that the limiting null distribution of the LRT for homogeneity is complex and cannot be approximated by the simpler chi-squared distribution [Hartigan, 1985; Chen and Chen, 2001]. To solve this problem, we used a parametric bootstrap to estimate the null distribution of the LRT based on the distribution parameters estimated from the observed data [McLachan, 1987].

As we have shown, analysis using phase-known data will have higher power to detect AE imbalance than using the phase-unknown data, particularly at low D'_{RT}. However, a variety of considerations can influence the choice to phase a given set of samples. Phase can be most accurately inferred for individuals with family data and investigators have chosen to use family based samples to maximize power to detect AE imbalance [Ge et al., 2009]. Alternatively, phase can be inferred in the absence of family information by the use of dense genotype data [Stephens et al., 2001; Stephens and Scheet, 2005; Marchini et al., 2006; Li et al., 2010]. If D'_{RT} is low and there are limited samples available for study, genotyping additional SNPs to locally phase haplotypes may substantially increase the power. However, genotyping SNPs could be cost-ineffective, particularly if only a single candidate rSNP is to be tested with a small number of genes in a region and the DNA and/or DNA samples are limited [see for example Fogarty et al., 2010]. In addition, in our simulations we assumed that the phasing of the data is accurate. However, phasing becomes less accurate with increasing distance between the rSNP and tSNP [Fallin and Schork, 2000] which will, in turn, decrease the power to detect AE imbalance with a phase-known test. In contrast, the F test is not affected by the rSNP-tSNP distance and may have higher power than the phase-known test to detect long-range *cis* effects.

When the rSNP and tSNP have similar allele frequencies and are in high D', our simulations show a decreased power for all the tests we proposed, due to smaller sample size for the rSNP homozygotes. In this situation, it may be useful to incorporate information from gDNA for all individuals. We attempted to apply an empirical Bayesian method [Mukherjee and Chatterjee, 2008; Chen et al., 2009] to improve the power by taking the weighted average of the AER means for the rSNP homozgotes cDNA and all individuals' gDNA. However this method could result in inflated type I error rate [Bhramar Mukherjee, personal communication] due to the potential difference in the AER means of the gDNA and cDNA data [see Fogarty et al., 2010].

We initially assumed a single rSNP influencing gene expression. To examine the sensitivity of the proposed tests to the presence of >1 rSNP, we studied the impact of an ungenotyped rSNP on the size and power of our tests to detect association between AE imbalance and the genotyped (putative) rSNP. We found that when the second ungenotyped rSNP is in LE with both the genotyped putative rSNP and the tSNP, the type I error rate of the tests is well controlled and that the power rankings of the various tests are essentially unchanged.

When the ungenotyped rSNP is in LD with the genotyped putative rSNP and the tSNP, we found that the 'false positive' rate of the tests can be high. In these instances the genotyped putative rSNP serves as a proxy for the ungenotyped rSNP and thus, when an association between AE imbalance and a potential rSNP is detected, we can at most infer that the AE imbalance is due to the putative rSNP and/or one or more other rSNP(s) in LD with the genotyped putative rSNP. Given this LD structure, the relative rankings of the tests remain essentially unchanged, while the absolute power of the tests can either decrease or increase depending on the frequencies of the expression-increasing allele of the two rSNPs and the direction of the effects of the two rSNPs. The test will have essentially no power to detect AE imbalance if the two rSNPs are in complete LD ($r^2=1$) and the effects of the two alleles on the same haplotype are of equal size but opposite directions. A second but unlikely scenario leading to no power is when the pairwise LD values for the three pairs of markers are all (near) zero, but the third-order LD is (near) one [Nielsen et al., 2004]. In this case, there are four three-locus haplotypes $\underline{R_G R_U T}$, $\underline{R_G r_U t}$, $\underline{r_G R_U t}$, and $\underline{r_G r_U T}$, each with probability ~.25, and correspondingly four haplogenotypes $\frac{R_G R_U T}{R_G r_U t}$, $\frac{R_G R_U T}{r_G R_U t}$, $\frac{r_G r_U T}{R_G r_U t}$, and $\frac{r_G r_U T}{r_G R_U t}$ also with probabilities ~.25. We did not observe a single example approaching this LD scenario in HapMap CEU samples on chromosome 1.

We have developed tests in the context of measurement of AER by SNP genotyping techniques that allow quantification of the AER of the two tSNP alleles. This work could be extended to RNA-seq data but would need to consider how to account for potential biases in mapping efficiency of the two tSNP alleles [Degner et al., 2009] and how to estimate the AER from sequence count data. Our proposed methods use the cDNA of rSNP homozygotes rather than the gDNA as reference for equal allelic expression. In the context of RNA-seq, using the rSNP homozygotes as the reference group has the advantage that it does not require high coverage gDNA sequencing.

In summary, in this paper we proposed three tests for association between AE imbalance and a *cis*-acting rSNP when phase is unknown and D' < 1 between the rSNP and a tSNP, and evaluated these tests plus existing tests for phase-unknown and phase-known data. We demonstrated that when AE imbalance is due to a single rSNP, the power of the tests is affected by multiple factors, including the LD between the rSNP and tSNP which has strong impact on the power ranking, and the allele frequencies of the two SNPs, number of tSNP heterozygotes, and AE imbalance effect size of the rSNP, which have less impact on the power ranking. We demonstrated that the presence of a second ungenotyped rSNP may reduce (or increase) statistical power, but seldom results in inconsistent tests, and tends not to modify the ranking of the tests. As general guidelines to maximize power to detect association between AE imbalance and a *cis*-acting rSNP, we recommend the use of the F test when the rSNP and tSNP are in or near LE (D'_{RT} ~0), the mixture-model based test when LD is intermediate ($.2 < D'_{RT} < .7$), and the t test when LD is high ($D'_{RT} > .7$).

Acknowledgements

We thank Drs. Michael Boehnke, Peter XK Song, and Bhramar Mukherjee for helpful discussions and Dr. Michael Boehnke for comments on the manuscript.

References

Bennett JH. One the theory of random mating. Ann Eugenics. 1954; 18:311–317.

Bray NJ, Jehu L, Moskvina V, Buxbaum JD, Dracheva S, Haroutunian V, Williams J, Buckland PR, Owen MJ, O'Donovan MC. Allelic expression of APOE in human brain: effects of epsilon status and promoter haplotypes. Hum Mol Genet. 2004; 13:2885–2892. [PubMed: 15385439]

Campino S, Forton J, Raj S, Mohr B, Auburn S, Fry A, Mangano VD, Vandiedonck C, Richardson A, Rockett K, Clark TG, Kwiatkowski DP. Validating discovered *cis*-acting regulatory genetic variants: application of an allele specific expression approach to HapMap populations. PLoS ONE. 2008; 3:e4105. [PubMed: 19116668]

Chen H, Chen J. Large sample distribution of the likelihood ratio test for normal mixtures. Canad J Statist. 2001; 29:201–216.

Chen Y-H, Chatterjee N, Carroll R. Shrinkage estimators for robust and efficient inference in haplotype-based case-control studies. JASA. 2009; 104:220–233. [PubMed: 19430598]

Cheung VG, Nayak RR, Wang IX, Elwyn S, Cousins SM, Morley M, Spielman RS. Polymorphic *cis*- and *trans*-regulation of human gene expression. PLoS Biol. 2010; 8:e1000480. [PubMed: 20856902]

Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, Burdick JT. Mapping determinants of human gene expression by regional and genome-wide association. Nature. 2005; 437:1365–1369. [PubMed: 16251966]

Cheung VG, Spielman RS. Genetics of human gene expression: mapping DNA variants that influence gene expression. Nat Rev Genet. 2009; 10:595–604. [PubMed: 19636342]

Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M. Mapping complex disease traits with global gene expression. Nat Rev Genet. 2009; 10:184–194. [PubMed: 19223927]

Cunnington MS, Santibánez Koref MF, Mayosi BM, Burn J, Keavney B. Chromosome 9p21 SNPs associated with multiple disease phenotypes correlate with ANRIL expression. PLoS Genet. 2010; 6:e1000899. [PubMed: 20386740]

Degner JF, Marioni JC, Pai AA, Pickrell JK, Nkador E, Gilad Y, Pritchard JK. Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. Bioinformatics. 2009; 25:3207–3212. [PubMed: 19808877]

Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. J Roy Statist Soc B. 1977; 39:1–38.

Fallin D, Schork NJ. Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-maximization algorithm for unphased diploid genotype data. Am J Hum Genet. 2000; 67:947–959. [PubMed: 10954684]

Ge B, Pokholok DK, Kwan T, Grundberg E, Morcos L, Verlaan DJ, Le J, Koka V, Lam KCL, Gagné, Dias J, Hoberman R, Montpetit A, Joly M-M, Harvey EJ, Sinnet D, Beaulieu P, Hamon R, Graziani A, Dewar K, Harmsen E, Majewski J, Goring HHH, Naumova AK, Blanchette M, Gunderson KL, Pastinen T. Global patterns of *cis* variation in human cells revealed by high-density allelic expression analysis. Nat Genet. 2009; 41:1216–1222. [PubMed: 19838192]

Gilad Y, Rifkin SA, Pritchard JK. Revealing the architecture of gene regulation: the promise of eQTL studies. Trends Genet. 2008; 24:408–415. [PubMed: 18597885]

Hartigan, JA. A failure of likelihood asymptotics for normal mixtures. In: LeCam, L.; Olshen, RA., editors. Proceedings of the Berk Conference in Honor of J. Neyman and J. Kiefer; 1985. p. 807-810.

Marchini J, Culter D, Patterson N, Stephens M, Eskin E, Lin S, Qin ZS, Munro HM, Abecasis GR, Donnelly P. A comparison of phasing algorithms for trios and unrelated individuals. Am J Hum Genet. 2006; 78:437–450. [PubMed: 16465620]

McLachlan GJ. On bootstrapping the likelihood ratio test statistic for the number of components in a normal mixture. Appl Statist. 1987; 36:318–324.

Monks SA, Leonardson A, Zhu H, Cundiff P, Pietrusiak P, Edwards S, Phillips JW, Sachs A, Schadt EE. Genetic inheritance of gene expression in human cell lines. Am J Hum Genet. 2004; 75:1094–1105. [PubMed: 15514893]

Mosteller F, Fisher RA. Combining independent tests of significance. Am Statist. 1948; 2:30–31.

Mukherjee B, Chatterjee N. Exploiting gene-environment independence for analysis of case-control studies: a shrinkage approach to trade off between bias and efficiency. Biometrics. 2008; 64:685–694. [PubMed: 18162111]

Nelder JA, Mead R. A simplex method for function minimization. Computer J. 1965; 7:308–313.

Nielson DM, Ehm MG, Zaykin DV, Weir BS. Effect of two- and three-locus linkage disequilibrium on the power to detect marker/phenotype associations. Genetics. 2004; 168:1029–1040. [PubMed: 15514073]

Pant PV, Tao H, Beilharz EJ, Ballinger DG, Cox DR, Frazer KA. Analysis of allelic differential expression in human white blood cells. Genome Res. 2006; 16:331–339. [PubMed: 16467561]

Pastinen T. Genome-wide allele-specific analysis: insights into regulatory variation. Nat Rev Genet. 2010; 11:533–538. [PubMed: 20567245]

Pastinen T, Sladek R, Gurd S, Sammak A, Ge B, Lepage P, Lavergne K, Villeneuve A, Gaudin T, Brandstrom H, Beck A, Verner A, Kingsley J, Harmsen E, Labuda D, Morgan K, Vohl MC, Naumova AK, Sinnett D, Hudson TJ. A survey of genetic and epigenetic variation affecting human gene expression. Physiol Genomics. 2003; 16:184–193. [PubMed: 14583597]

Pastinen T, Ge B, Hudson TJ. Influence of human genome polymorphism on gene expression. Hum Mol Genet. 2006; 15:R9–R16. [PubMed: 16651375]

Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, Veyrieras JB, Stephens M, Gilad Y, Pritchard JK. Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature. 2010; 464:768–772. [PubMed: 20220758]

Santibánez Koref MF, Wilson V, Cartwright N, Cunnington MS, Mathers JC, Bishop DT, Curtis A, Dunlop MG, Burn J. MLH1 differential allelic expression in mutation carriers and controls. Ann Hum Genet. 2010; 74:479–488. [PubMed: 20860725]

Serre D, Gurd S, Ge B, Sladek R, Sinnett D, Harmsen E, Bibikova M, Chudin E, Barker DL, Dickinson T, Fan J-B, Hudson TJ. Differential allelic expression in the human genome: a robust approach to identify genetic and epigenetic *cis*-acting mechanisms regulating gene expression. PLoS Genet. 2008; 4:e1000006. [PubMed: 18454203]

Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, Allen HL, Lindgren CM, Luan J, Mägi R, Randall JC, Vedantam S, Winkler TW, Qi L, Workalemahu T, Heid IM, Steinthorsdottir V, Stringham HM, Weedon MN, Wheeler E, Wood AR, Ferreira T, Weyant RJ, Segrè AV, Estrada K, Liang L, Nemesh J, Park JH, Gustafsson S, Kilpeläinen TO, Yang J, Bouatia-Naji N, Esko T, Feitosa MF, Kutalik Z, Mangino M, Raychaudhuri S, Scherag A, Smith AV, Welch R, Zhao JH, Aben KK, Absher DM, Amin N, Dixon AL, Fisher E, Glazer NL, Goddard ME, Heard-Costa NL, Hoesel V, Hottenga JJ, Johansson A, Johnson T, Ketkar S, Lamina C, Li S, Moffatt MF, Myers RH, Narisu N, Perry JR, Peters MJ, Preuss M, Ripatti S, Rivadeneira F, Sandholt C, Scott LJ, Timpson NJ, Tyrer JP, van Wingerden S, Watanabe RM, White CC, Wiklund F, Barlassina C, Chasman DI, Cooper MN, Jansson JO, Lawrence RW, Pellikka N, Prokopenko I, Shi J, Thiering E, Alavere H, Alibrandi MT, Almgren P, Arnold AM, Aspelund T, Atwood LD, Balkau B, Balmforth AJ, Bennett AJ, Ben-Shlomo Y, Bergman RN, Bergmann S, Biebermann H, Blakemore AI, Boes T, Bonnycastle LL, Bornstein SR, Brown MJ, Buchanan TA, Busonero F, Campbell H, Cappuccio FP, Cavalcanti-Proença C, Chen YD, Chen CM, Chines PS, Clarke R, Coin L, Connell J, Day IN, Heijer M, Duan J, Ebrahim S, Elliott P, Elosua R, Eiriksdottir G, Erdos MR, Eriksson JG, Facheris MF, Felix SB, Fischer-Posovszky P, Folsom AR, Friedrich N, Freimer NB, Fu M, Gaget S, Gejman PV, Geus EJ, Gieger C, Gjesing AP, Goel A, Goyette P, Grallert H, Grässler J, Greenawalt DM, Groves CJ, Gudnason V, Guiducci C, Hartikainen AL, Hassanali N, Hall AS, Havulinna AS, Hayward C, Heath AC, Hengstenberg C, Hicks AA, Hinney A, Hofman A, Homuth G, Hui J, Igl W, Iribarren C, Isomaa B, Jacobs KB, Jarick I, Jewell E, John U, Jørgensen T, Jousilahti P, Jula A, Kaakinen M, Kajantie E, Kaplan LM, Kathiresan S, Kettunen J, Kinnunen L, Knowles JW, Kolcic I, König IR, Koskinen S, Kovacs P, Kuusisto J, Kraft P, Kvaløy K, Laitinen J, Lantieri O, Lanzani C, Launer LJ, Lecoeur C, Lehtimäki T, Lettre G, Liu J, Lokki ML, Lorentzon M, Luben RN, Ludwig B, MAGIC; Manunta P, Marek D, Marre M, Martin NG, McArdle WL, McCarthy A, McKnight B, Meitinger T, Melander O, Meyre D, Midthjell K, Montgomery GW, Morken MA, Morris AP, Mulic R, Ngwa JS, Nelis M, Neville MJ, Nyholt DR, O'Donnell CJ, O'Rahilly S, Ong KK, Oostra B, Paré G, Parker AN, Perola M, Pichler I, Pietiläinen KH, Platou CG, Polasek O, Pouta A, Rafelt S, Raitakari O, Rayner NW, Ridderstråle M, Rief W, Ruokonen A, Robertson NR, Rzehak P, Salomaa V, Sanders AR, Sandhu

MS, Sanna S, Saramies J, Savolainen MJ, Scherag S, Schipf S, Schreiber S, Schunkert H, Silander K, Sinisalo J, Siscovick DS, Smit JH, Soranzo N, Sovio U, Stephens J, Surakka I, Swift AJ, Tammesoo ML, Tardif JC, Teder-Laving M, Teslovich TM, Thompson JR, Thomson B, Tönjes A, Tuomi T, van Meurs JB, van Ommen GJ, Vatin V, Viikari J, Visvikis-Siest S, Vitart V, Vogel CI, Voight BF, Waite LL, Wallaschofski H, Walters GB, Widen E, Wiegand S, Wild SH, Willemsen G, Witte DR, Witteman JC, Xu J, Zhang Q, Zgaga L, Ziegler A, Zitting P, Beilby JP, Farooqi IS, Hebebrand J, Huikuri HV, James AL, Kähönen M, Levinson DF, Macciardi F, Nieminen MS, Ohlsson C, Palmer LJ, Ridker PM, Stumvoll M, Beckmann JS, Boeing H, Boerwinkle E, Boomsma DI, Caulfield MJ, Chanock SJ, Collins FS, Cupples LA, Smith GD, Erdmann J, Froguel P, Grönberg H, Gyllensten U, Hall P, Hansen T, Harris TB, Hattersley AT, Hayes RB, Heinrich J, Hu FB, Hveem K, Illig T, Jarvelin MR, Kaprio J, Karpe F, Khaw KT, Kiemeney LA, Krude H, Laakso M, Lawlor DA, Metspalu A, Munroe PB, Ouwehand WH, Pedersen O, Penninx BW, Peters A, Pramstaller PP, Quertermous T, Reinehr T, Rissanen A, Rudan I, Samani NJ, Schwarz PE, Shuldiner AR, Spector TD, Tuomilehto J, Uda M, Uitterlinden A, Valle TT, Wabitsch M, Waeber G, Wareham NJ, Watkins H, Procardis Consortium. Wilson JF, Wright AF, Zillikens MC, Chatterjee N, McCarroll SA, Purcell S, Schadt EE, Visscher PM, Assimes TL, Borecki IB, Deloukas P, Fox CS, Groop LC, Haritunians T, Hunter DJ, Kaplan RC, Mohlke KL, O'Connell JR, Peltonen L, Schlessinger D, Strachan DP, van Duijn CM, Wichmann HE, Frayling TM, Thorsteinsdottir U, Abecasis GR, Barroso I, Boehnke M, Stefansson K, North KE, McCarthy MI, Hirschhorn JN, Ingelsson E, Loos RJ. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet. 2010; 42:937–948. [PubMed: 20935630]

Stephens M, Scheet P. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet. 2005; 76:449–462. [PubMed: 15700229]

Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet. 2001; 68:978–989. [PubMed: 11254454]

Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, Lyle R, Hunt S, Kahl B, Antonarakis SE, Tavaré S, Deloukas P, Dermitzakis ET. Genome-wide associations of gene expression variation in humans. PLoS Genet . 2005; 6:695–704.

Tao H, Cox DR, Frazer KA. Allele-specific *KRT1* expression is a complex trait. PLoS Genet. 2006; 2:e93. [PubMed: 16789827]

Teare MD, Heighway J, Santibánez Koref MF. An exprectation-maximization algorithm for the analysis of allelic expression imbalance. Am J Hum Genet. 2006; 79:539–543. [PubMed: 16909391]

Thomson G, Baur MP. Third order linkage disequilibrium. Tissue Antigens. 24:250–255. [PubMed: 6515638]

Yan H, Yuan W, Velculescu VE, Vogelstein B, Kinzler KW. Allelic Variation in Human Gene Expression. Science. 2002; 297:1143. [PubMed: 12183620]

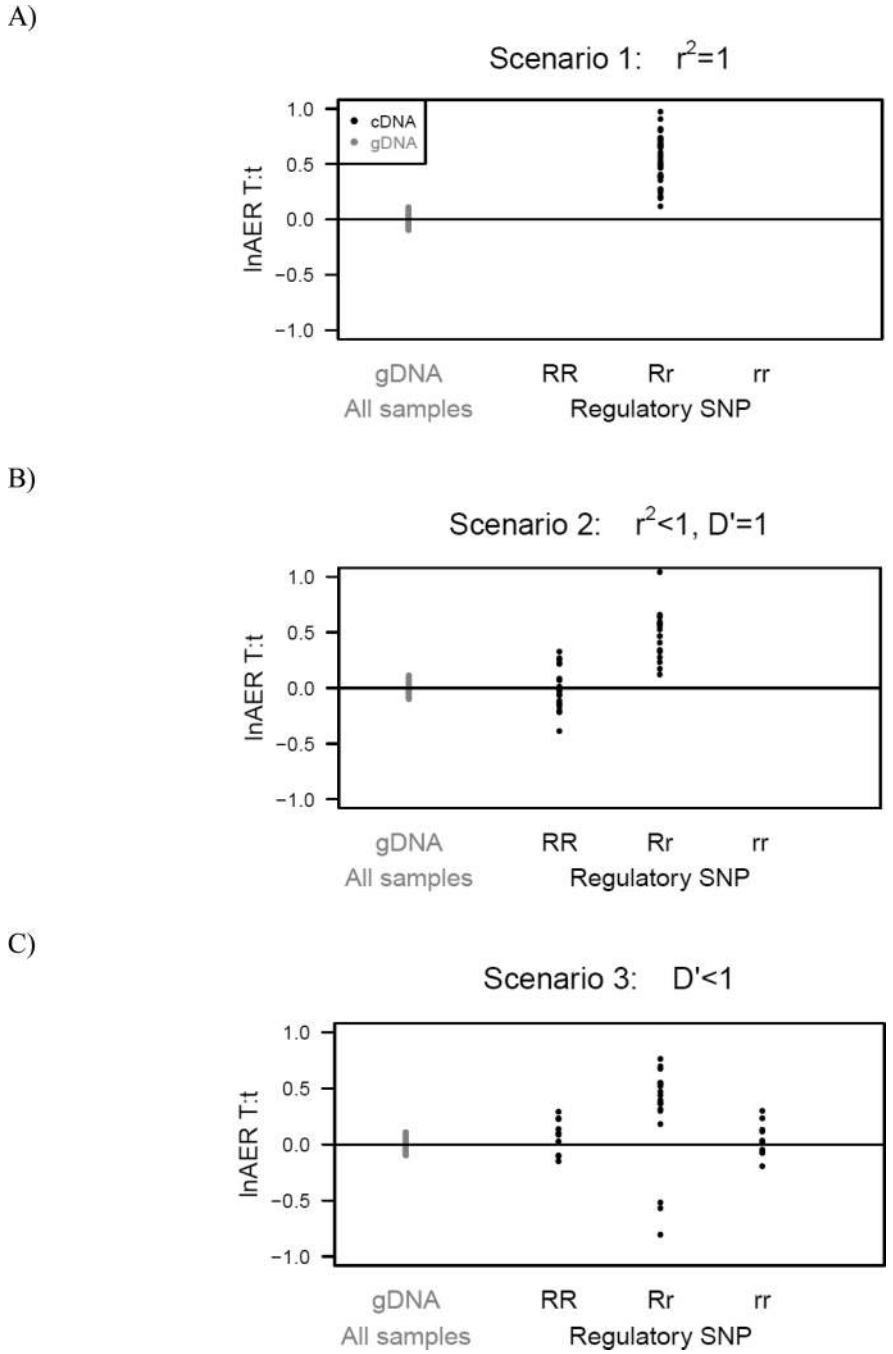


Figure 1.

The lnAER data patterns for three different LD structures between the rSNP and tSNP.

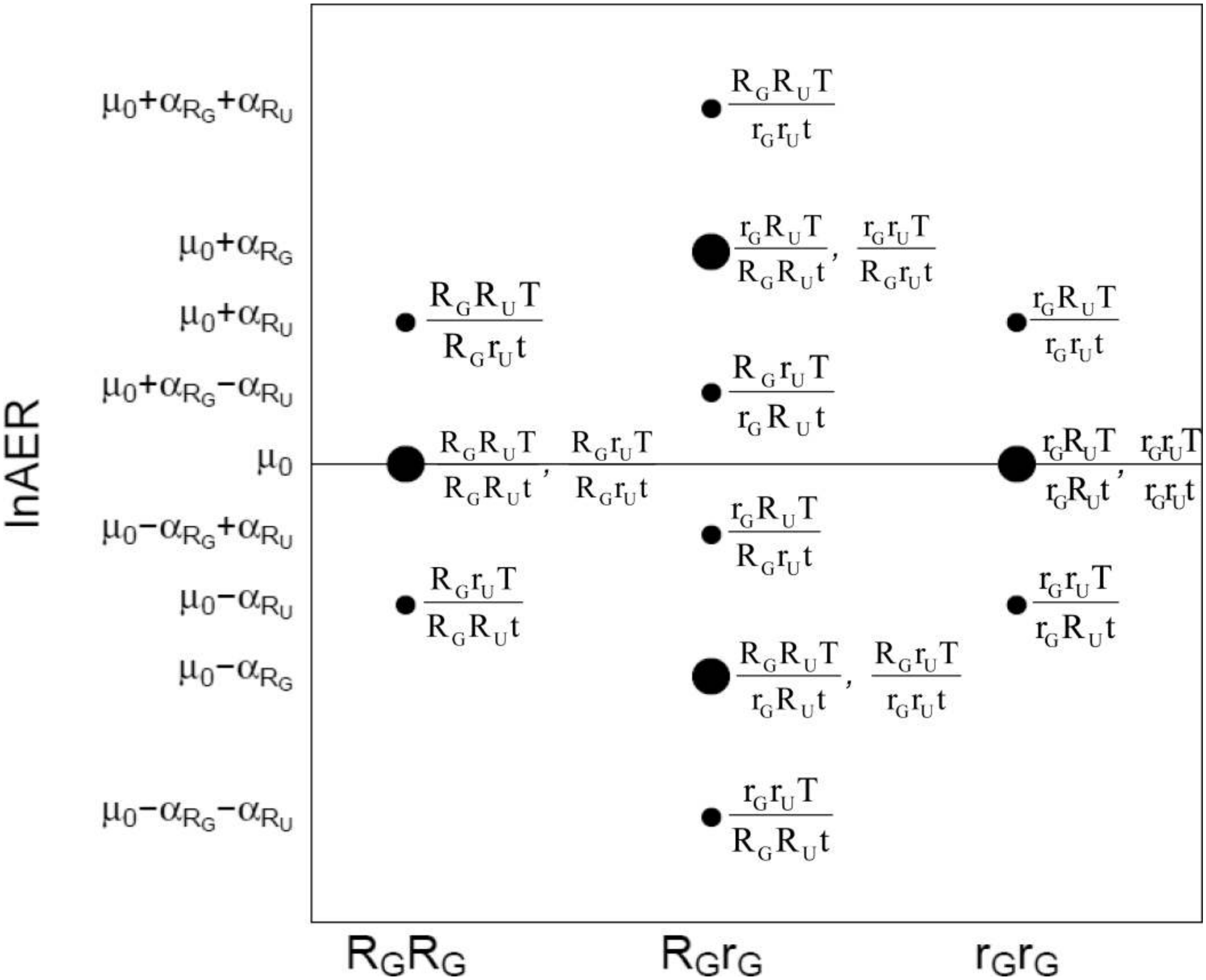


Figure 2.

The expected lnAER data pattern when there is a second ungenotyped rSNP. In this example, the allele frequencies for the two rSNPs and the tSNP are equal $p_{R_G} = p_{R_U} = p_T = .5$, and the three loci are independent. Assume the effect of the genotyped rSNP on lnAER is greater than that of the ungenotyped rSNP ($\alpha_{R_G} > \alpha_{R_U}$) and the two rSNPs act additively. Position and size of each circle represent the mean lnAER and the frequency of the corresponding haplogenotype(s) to its right, respectively.

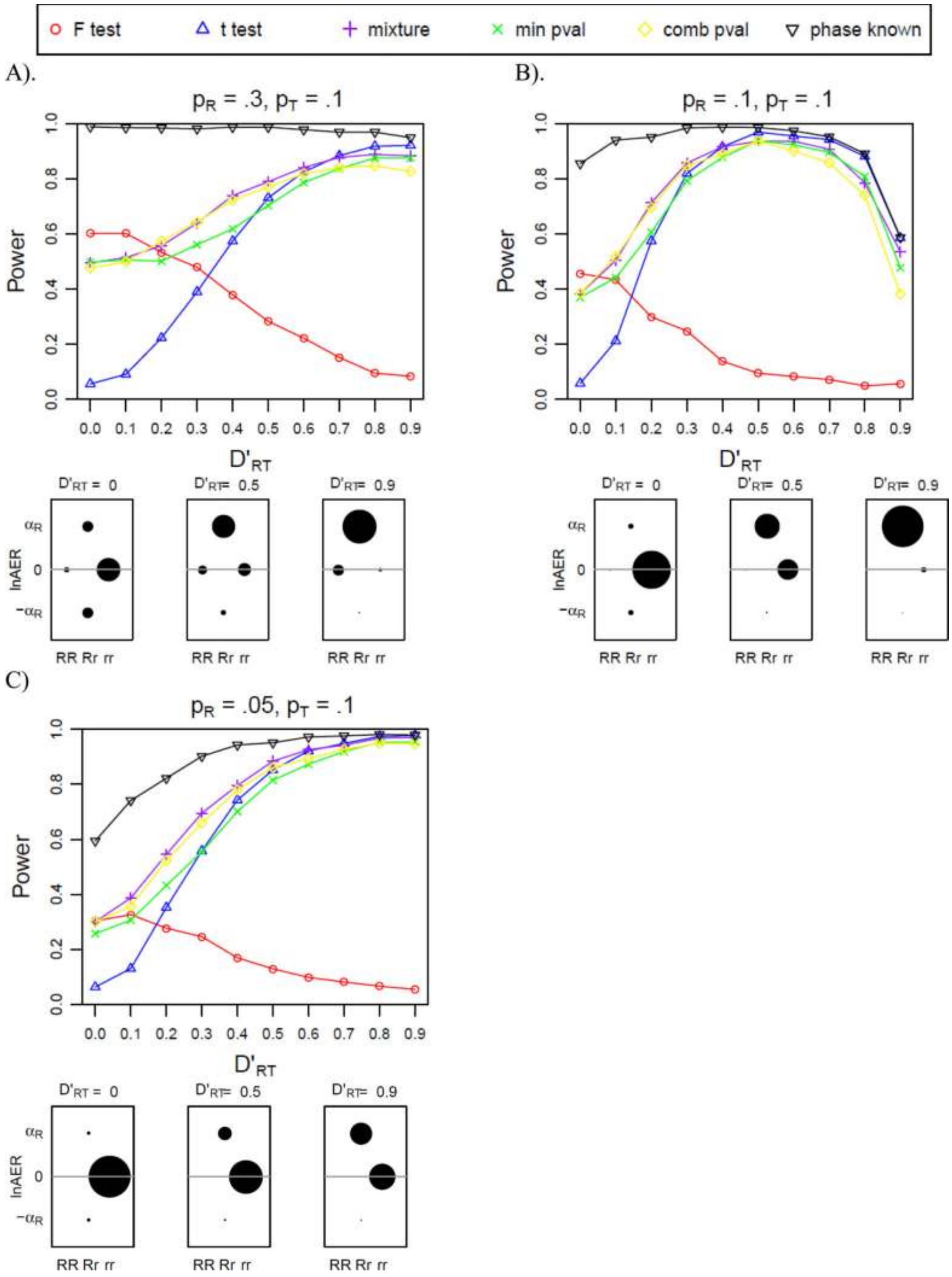
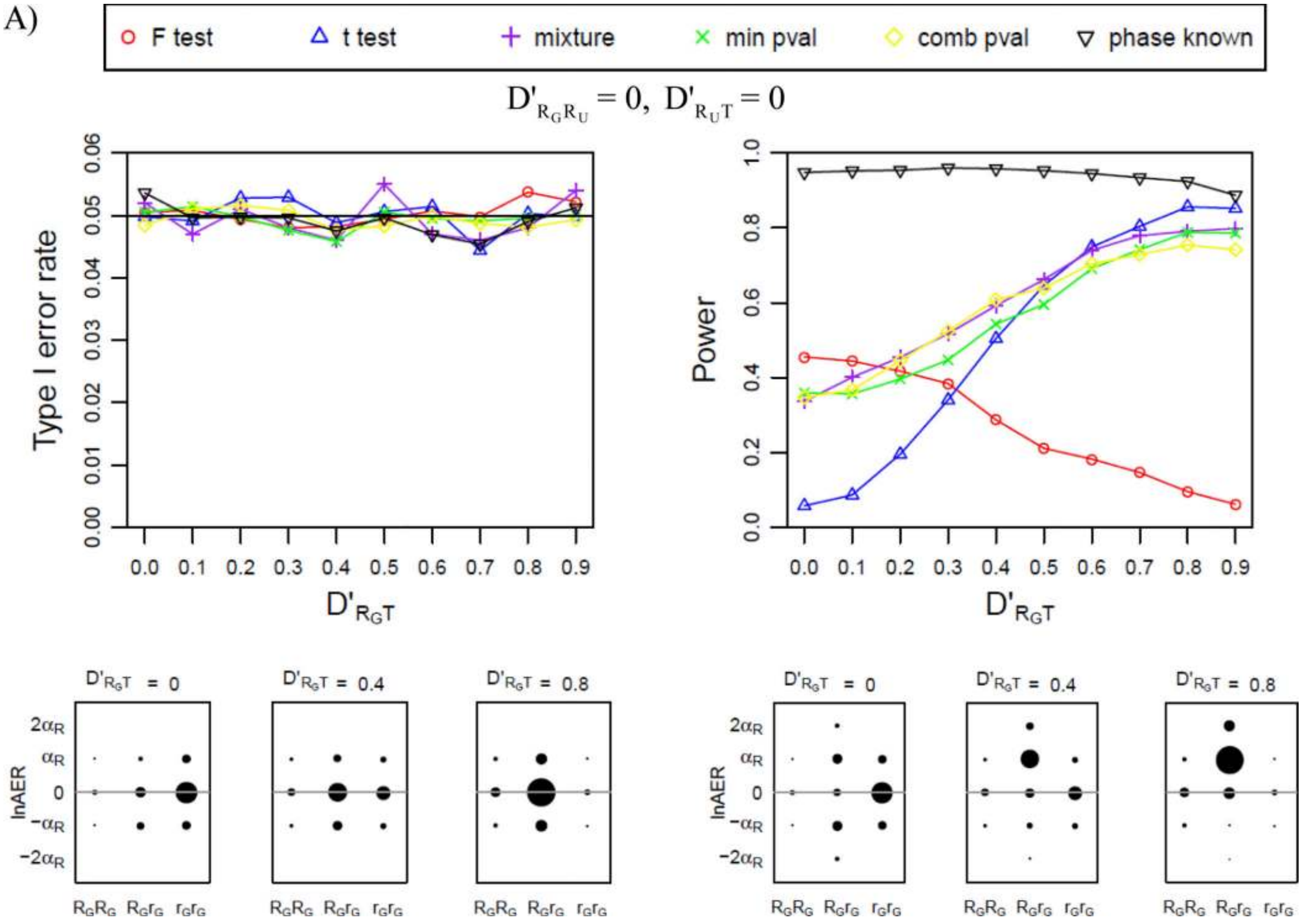


Figure 3.

Power of the tests at significance level α = .05 when the number of tSNP heterozygotes *N* = 100, AE imbalance effect size α_R = .85 with variance σ^2 = 1, and allele frequency of the tSNP is p_T = .1 and of the rSNP is A) p_R =.3, B) p_R =.1 and C) p_R =.05.

Balloon plot under each panel is the expected lnAER pattern under different D'$_{RT}$ values between the rSNP and tSNP. The position and diameter of each dot represent the mean lnAER and the frequency of the corresponding haplogenotype, respectively.

P-values are estimated using 1000 permutations for the F, t, minimum-p-value and combined-p-value tests, and 1000 bootstraps for the mixture-model based test; power for each test is calculated based on 1000 simulation replicates.


A)
○ F test
△ t test
+ mixture
× min pval
◇ comb pval
▽ phase known
$D'_{R_G R_U} = 0,\ D'_{R_U T} = 0$
Type I error rate
0.00
0.01
0.02
0.03
0.04
0.05
0.06
Power
0.0
0.2
0.4
0.6
0.8
1.0
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
$D'_{R_G T}$
$D'_{R_G T} = 0$
$D'_{R_G T} = 0.4$
$D'_{R_G T} = 0.8$
lnAER
$2\alpha_R$
α_R
0
$-\alpha_R$
$-2\alpha_R$
R_GR_G R_Gr_G r_Gr_G

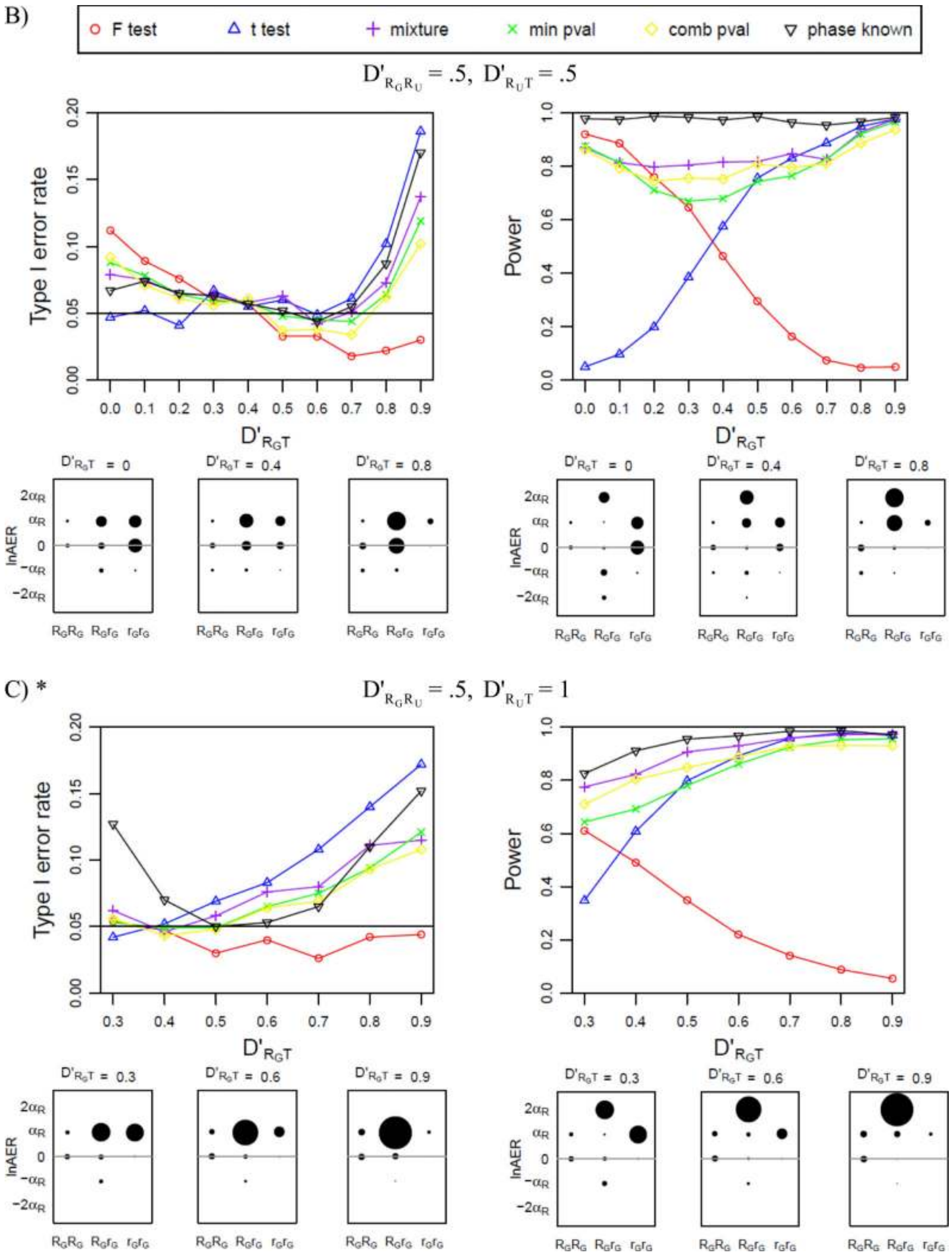


Figure 4.

Impact of a second ungenotyped rSNP (R_U) on type I error rate (left panel) and power (right panel) of the tests to detect association between AE imbalance and the genotyped rSNP (R_G) at significance level $\alpha = .05$ under different LD structures: A) $D'_{R_GR_U} = 0$, $D'_{R_UT} = 0$; B) $D'_{R_GR_U} = .5$, $D'_{R_UT} = .5$; and C) $D'_{R_GR_U} = .5$, $D'_{R_UT} = 1$. For all plots, the third order LD $D'_{R_GR_UT} = 0$, $N = 100$ tSNP heterozygotes with MAF $p_T = .1$. Allele frequencies for the genotyped and ungenotyped rSNPs $p_{R_G} = p_{R_U} = .3$. For the type I error rate estimation (left panel), the effect size of the genotyped and ungenotyped rSNPs on lnAER are $\alpha_{R_G} = 0$ and

α_{R_U} =.85 with variance $\sigma^2 = 1$, respectively. For the power estimation (right panel), the genotyped and ungenotyped rSNPs act additively and have equal effect size on lnAER $\alpha_{R_G} = \alpha_{R_U} = .85$, each with variance $\sigma^2 = 1$.

P-values are estimated using 1000 permutations for the F, t, minimum-p-value and combined-p-value tests, and 1000 bootstraps for the mixture-model based test; type I error and power for each test are calculated based on 1000 simulation replicates.

* : $D'_{R_G T}$ cannot go below .3 given the allele frequencies and the LD structure of the three SNPs.

Table 1

At a given D'_{RT}, the number of tSNP heterozygotes *N* needed for the phase-unknown tests to obtain similar power level as for the phase-known test. Allele frequencies of the rSNP and tSNP are $p_R = .3$, $p_T = .1$. Significance level $\alpha = .05$.

***N* for phase-known test (α_R)**	**D'_{RT}**	**Power for phase-known test**	***N* for phase-unknown tests**		
			F	**Mixture**	**t**
50 (1.2)	0	.99	125	150	>500
	.3	.98	190	135	340
	.6	.97	>500	95	100
	.9	.93	>500	70	60
100 (.85)	0	.99	400	460	>1000
	.3	.98	580	300	600
	.6	.97	>1000	175	185
	.9	.94	>1000	115	110