

A Graphical Weighted Power Improving Multiplicity Correction Approach for SNP Selections

Garrett Saunders¹, Guifang Fu^{1,*} and John R. Stevens¹

¹Department of Mathematics and Statistics, Utah State University, Logan, UT 84322, USA

Abstract: Controlling for the multiplicity effect is an essential part of determining statistical significance in large-scale single-locus association genome scans on Single Nucleotide Polymorphisms (SNPs). Bonferroni adjustment is a commonly used approach due to its simplicity, but is conservative and has low power for large-scale tests. The permutation test, which is a powerful and popular tool, is computationally expensive and may mislead in the presence of family structure. We propose a computationally efficient and powerful multiple testing correction approach for Linkage Disequilibrium (LD) based Quantitative Trait Loci (QTL) mapping on the basis of graphical weighted-Bonferroni methods. The proposed multiplicity adjustment method synthesizes weighted Bonferroni-based closed testing procedures into a powerful and versatile graphical approach. By tailoring different priorities for the two hypothesis tests involved in LD based QTL mapping, we are able to increase power and maintain computational efficiency and conceptual simplicity. The proposed approach enables strong control of the familywise error rate (FWER). The performance of the proposed approach as compared to the standard Bonferroni correction is illustrated by simulation and real data. We observe a consistent and moderate increase in power under all simulated circumstances, among different sample sizes, heritabilities, and number of SNPs. We also applied the proposed method to a real outbred mouse HDL cholesterol QTL mapping project where we detected the significant QTLs that were highlighted in the literature, while still ensuring strong control of the FWER.

Received on: July 27, 2014 - Revised on: September 08, 2014 - Accepted on: September 08, 2014

Keywords: Graphical weighted approach, Multiple correction, Hypothesis testing, QTL, Linkage disequilibrium.

INTRODUCTION

Linkage Disequilibrium (LD) analysis plays a fundamental role in gene mapping, as a tool for uncovering biological trait regulating genes. Many biological traits are influenced by genetic variants and hence it is possible to determine the rough genomic position of the causative variations through associations between SNPs and phenotype [1-13]. Among the popular SNP selection approaches, the single-locus association with solid multiplicity correction ability remains a powerful tool as associations can generally only be found over small distances [14]. Moreover, as tens-of-thousands of SNPs for genome-wide association studies (GWAS) are under demand [5], the single SNP based LD analysis can at least provide a necessary initial screening selection to detect a subset of promising candidates for further exploration [15, 16].

Despite the great progress which has already been made within LD based Quantitative Trait Loci (QTL) mapping, powerful and computationally efficient methods for large-scale simultaneous testing of individual SNPs with strong control of the familywise error rate (FWER) are still lacking [17, 18]. FWER is the most accepted error rate used to determine significance levels for large-scale testing problems

where the goal is to provide conclusive results. In some studies, researchers often control the False Discovery Rate (FDR) to obtain a large pool of potentially significant SNPs and then select only the most significant subset for validation due to cost restrictions. However, this rule can lead to unwanted results as the FDR is controlled only for *all* selected SNPs, and provides no promise of control for an arbitrarily selected subset of the significant SNPs. Thus, we recommend controlling the FWER (in place of the FDR) in exploratory scenarios where only the most promising results will be considered.

The Bonferroni correction, as one of the most widely used statistical procedures, is often employed to control the FWER when multiple tests are conducted. However, the Bonferroni correction is not favorable in large-scale testing because it substantially reduces the statistical power, hence decreasing the chances of detecting SNPs with real effects [19]. While permutation procedure has been widely employed to adjust for correlated tests, it is computationally expensive [15, 20, 21] and may mislead in the presence of family structure [22]. Moreover, the permutation approach was designed for only one test setting. In LD based QTL studies, the high likelihood of dependencies among SNPs and the two tests structure strongly demand a new multiplicity adjustment approach that maintains simplicity but is more powerful.

In the LD based QTL model [13], detecting a significant QTL that is associated with a certain phenotype requires two

*Address correspondence to this author at the Department of Mathematics and Statistics, Utah State University, Logan, UT 84322, USA; Tel: +001 435 797 0749; E-mail: Guifang.fu@usu.edu.

hypothesis tests, one testing for the existence of a QTL for a given SNP (i.e. whether or not the QTL is associated with the phenotype), and the other testing for the strength of the LD between the SNP and the existing QTL (i.e. whether or not the QTL is successfully detected by the model). Although the existence of a significant QTL is the ultimate goal, the degree of LD between the QTL and SNP is also critical in guaranteeing the basic assumptions of the model. By Fu *et al.*'s assumptions [13], the unobservable QTL can be mapped by its association with the observable SNP through the conditional probability of the genotype of the QTL given the genotype of the SNP. Therefore, only QTLs that are not only significantly existing but also strongly linked with a SNP will be identified, i.e. identifying a significant QTL requires rejecting these two hypotheses simultaneously. Although the LD based QTL model has been successful in locating significant QTLs [13, 23, 24], the Bonferroni multiplicity correction approach used previously ignored two important issues. First, if the QTL existence test fails to reject, then the LD between the SNP and QTL is not identifiable in their mixture model. Second, the Bonferroni correction is too conservative for large-scale of SNPs.

In this article, we propose a new power improving multiplicity correction approach specially designed for the LD based QTL mapping, on the basis of graphical weighted-Bonferroni methods [25]. By introducing a logical structuring for the two tests involved for each SNP, i.e. higher order for QTL existence testing (primary) than the LD testing (secondary), the LD test will never be investigated if the primary test concludes that the QTL does not exist. By exploiting the priority ordering of the two hypotheses to adjust the p-values, our proposed approach can avoid the previously mentioned identifiability issue, and address the multiplicity correction for large-scale number of SNPs. None of the current LD based QTL methodologies directly overcome these challenges when performing these two tests [13, 23, 24]. Our proposed multiple correction approach with priority structuring has been shown to synthesize weighted Bonferroni-based closed testing procedures such as the weighted Bonferroni-Holm procedure, fixed sequence tests, gatekeeping procedures, and the fallback procedure into a powerful and versatile graphical approach [25], which we tailor here for the LD based QTL mapping.

In the following section we present the LD based QTL model and the two tests involved. Next, we describe in detail how we design the logical structuring to perform the multiplicity correction for the LD based QTL model. The significance of the power advantage of the proposed method over the Bonferroni correction is established through both simulations and one real QTL mapping project. Since sample size, heritability, and number of SNPs all determine the power of the method, we illustrate the power through heritability of 0.1 and 0.4, sample size small (100), medium (300), and large (500), and number of SNPs changing from 1, 10, 50, 100, 500, to 1,000. We end with a discussion of the results.

METHODOLOGY

LD Based QTL Mapping Model

To map the rough location of the QTL regulating a certain biological trait, we apply the mixture model [13]. Under this model, QTL is detected by statistically modeling

the genotypic variation through not only the association between phenotype and the putative QTL, but also the association between the putative QTL and SNP. Since the SNP genotype is observable, we can infer the probabilities of a putative QTL genotype by the conditional probability of QTL (*A*) genotype given the SNP (*M*) genotype, as long as there exists LD between the SNP and putative QTL [26].

The mixture model of [13] assumes each individual's phenotype $Y_i, i = 1, \dots, n$, is a random variable from density $f_l(Y_i | \theta_l)$, where $l \in \{1, 2, 3\}$ denotes three distinct QTL genotypes. Each QTL genotype is assumed to induce a separate distribution of phenotypes. Typically, normal distributions are assumed for each $f_l(Y_i | \theta_l)$ with $\theta_l = (\mu_l, \sigma)$. From these assumptions, the corresponding likelihood is expressed as [13]

$$L(\mu_l, \sigma | Y, M) = \prod_{i=1}^n \sum_{l=1}^3 \omega_{li} f_l(Y_i | \mu_l, \sigma), \tag{1}$$

where ω_{li} is the conditional probability of individual *i* having QTL genotype *l* given their SNP genotypes, μ_l is the phenotypic mean for QTL genotype *l*, σ is the common standard deviation for all genotypes, and $f_l(Y_i | \mu_l, \sigma)$ is the probability density of observations for individual *i* at QTL genotype *l* [13, 26, 28].

The probability of the SNP's major allele (*M*) is denoted by *p*, and correspondingly $1 - p$ for the minor allele (*m*). Similarly, the probability of the QTL's major allele (*A*) is denoted by *q*, and correspondingly $1 - q$ for the minor allele (*a*). Together, the SNP and QTL form four haplotypes (*MA, Ma, mA, and ma*) with corresponding frequencies $p_{11} = pq + D, p_{10} = p(1 - q) - D, p_{01} = (1 - p)q - D,$ and $p_{00} = (1 - p)(1 - q) + D,$ respectively. Here, *D* is the linkage disequilibrium between SNP and QTL. The conditional probabilities ω_{li} of the QTL's various genotypes (*AA, Aa, and aa*) can be calculated upon the observed SNP genotypes (*MM, Mm, and mm*) from the joint probabilities [13, 27]. Hence, ω_{li} is a function of *p, q,* and *D*. The EM algorithm is then applied to the likelihood in (1) to obtain maximum likelihood estimates for all parameters [13, 27].

Two Hypothesis Tests

Through the likelihood in (1), the hypotheses

$$H_0^L : \mu_1 = \mu_2 = \mu_3 \quad \text{vs} \\ H_1^L : \text{one of the equalities above does not hold} \tag{2}$$

can be used to test if the QTL is significantly associated with phenotype *Y* (i.e. existence of QTL). Since all the unknown parameters in (1) were estimated by maximum likelihood estimates (MLEs), a log likelihood ratio statistic can be used to test the hypotheses in (2) [13]. The resulting test statistic (χ_L^2) is asymptotically distributed as a χ^2 under H_0^L for large enough samples [13, 26].

On the other hand, linkage disequilibrium, denoted by D , between the SNP and QTL can be tested by means of the hypotheses

$$H_0^D : D = 0 \quad \text{vs} \quad H_1^D : D \neq 0. \quad (3)$$

Once the existence of a QTL is established, the test statistic used to judge whether or not the QTL is significantly associated with SNP is [13, 29]:

$$\chi_D^{2*} = \frac{2n\hat{D}^2}{\hat{p}(1-\hat{p})\hat{q}(1-\hat{q})} \quad (4)$$

$$= 2nr^2. \quad (5)$$

Here, r^2 is the square of the correlation coefficient between the SNP and QTL that has been used in most of the related literature, which has many good sampling properties [30, 31]. Under H_0^D , χ_D^{2*} is asymptotically distributed as χ_1^2 , from which the tail probability (p-value) of the observed level of association can be determined [10, 23, 26, 29, 32].

Because whether or not a QTL exists and whether or not the existing QTL is able to be successfully identified by the model are both critical, the QTL will not be interpreted as significant unless the two hypothesis tests (2) and (3) are both rejected. Three problems urge the improvement on this two hypothesis tests setting. Firstly, D is not identifiable under the null hypothesis $H_0^L : \mu_1 = \mu_2 = \mu_3 = \mu$ [33, 34]. That is, the parameter ω_{l_i} falls out of the model when the means are equal, as the $f_l(Y_i | \mu_l, \sigma), l=1,2,3$ in the likelihood (1) are identical in this case, resulting in $\sum_{i=1}^3 \omega_{l_i} f_l(Y_i | \mu_l, \sigma) = f(Y_i | \mu, \sigma)$. Hence, the likelihood reduces to $L(\mu, \sigma | Y, M) = \prod_{i=1}^n f(Y_i | \mu, \sigma)$, so that D which is computed from ω_{l_i} can not be computed. As a result, testing $D \neq 0$ under H_0^L is not meaningful. Secondly, a multiplicity correction is needed for simultaneous testing of all SNPs. Thirdly, the existence of a QTL underlying certain biological traits is the ultimate goal for the real application. Inspired by the idea of the graphical Bonferroni approach [25], we set the existence of QTL (2) to be the primary test and the LD test (3) to be the secondary test. If the primary test is not rejected, the secondary test will not be investigated. As a result, our proposed multiplicity correction approach increases the power, while preserving strong control of FWER and avoiding the identifiability issue of D under H_0^L .

Graphical Bonferroni Approach

The Graphical Bonferroni Approach (GBA) is a versatile and easily communicated general adjustment method for multiple testing [25]. Provided as a generalized framework, it must be specially tailored to each testing situation. Generally speaking, it is most powerful for situations where hypotheses can be partitioned into levels of importance such that the most important hypotheses are tested first and the lower level hypotheses are tested only if the higher level hypotheses show significant results.

All hypotheses of interest are depicted as nodes in a directed acyclic graph. Local significance thresholds for each node (hypothesis) dictate the local level at which each hypothesis is tested. Weighted edges between all nodes map the logical structuring of the designated testing approach. When a hypothesis is rejected, the weighted edges dictate the proportion of the locally assigned significance threshold that is passed from the rejected node to all connected nodes. Thus, the graph induces an iterative testing approach that is shown to result in a closed-test that admits a short-cut [25]. Further, Algorithm 1 of [25] provides a simple updating technique that performs the short-cut. Strong control of the FWER at level α is proven to occur so long as three regularity conditions are met: 1) the sum of the local significance thresholds is no more than α , 2) the sum of outgoing edge weights from each node are no larger than unity, and 3) no node has an edge connecting to itself [25].

Rejection Scheme

Since the ultimate goal is to check existence of QTL, the first interest is in testing H_0^L in (2) to see if the phenotype shows evidence of association with a latent QTL. Depending on the results of the test of (2), the testing for the given SNP will either end, or interest will be turned to testing H_0^D in (3) to see if the SNP is also associated with the QTL. (Fig. 1A) demonstrates how all of α is used to test the first hypothesis, H_0^L , and none of α is initially given to the testing of H_0^D . That is, node H_0^L has local significance threshold α , and H_0^D has local significance threshold 0. Assuming H_0^L is claimed significant, the node belonging to H_0^L would be removed and all of α passed on to H_0^D as signified by the edge weight of 1 along the path from H_0^L to H_0^D . At this point, H_0^D is tested at level α , its new local significance threshold given the rejection of H_0^L as shown in (Fig. 1B).

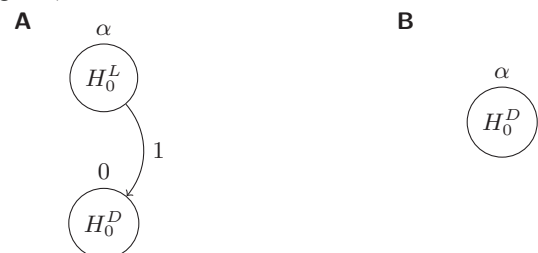


Fig. (1). **A:** Demonstration of the GBA testing scheme for a single SNP. **B:** The updated graph after finding H_0^L significant.

Note that adjusted p -values could be similarly obtained for each node. The adjusted p -value for H_0^L would be the same as the unadjusted value. The adjusted p -value for H_0^D would be either larger of its unadjusted value and H_0^L 's unadjusted value (if H_0^L was significant at level α) or 1 (if

H_0^L was not significant at level α). The structuring ensures that a child node (such as H_0^D in Fig. 1) cannot have a smaller adjusted p -value than its parent node (H_0^L in Fig. 1).

As a result, for the single SNP analysis, either both hypotheses will be tested at level α , or the testing will stop after H_0^L without considering H_0^D . Alternatively, the Bonferroni correction would test both hypotheses at $\alpha/2$. Hence, the Bonferroni adjustment has less power, due to its smaller thresholds. Compared to the traditional Bonferroni, the only potential disadvantage of the GBA method is that it skips testing H_0^D if H_0^L is not significant. However, this potential disadvantage becomes an advantage for the LD based QTL model because D is not identifiable under the null of H_0^L . Thus, the only situation in which the Bonferroni method would have a possible advantage over the GBA method is not applicable here.

Our proposed GBA method further achieves an advantage in the case of multiple SNPs through sharing of the α -level between SNPs. Say there are m SNPs to be tested for both H_0^L and H_0^D . Let H_0^{Li} and H_0^{Di} denote these two hypotheses respectively for the i th SNP, $i = 1, \dots, m$. The GBA approach begins by allocating initial local thresholds of α/m to each $H_0^{Li}, i = 1, \dots, m$. Initial local thresholds of zero are allocated to each $H_0^{Di}, i = 1, \dots, m$. Edge weights of 1 are established from each H_0^{Li} to the corresponding H_0^{Di} for all $i = 1, \dots, m$. Edges with weight $1/(m-1)$ are also established between each H_0^{Di} and all H_0^{Lj} for each $i = 1, \dots, m$ and $j \neq i$. Multiple testing (or multiplicity adjustment) then proceeds according to Algorithm 1 of [25].

Fig. 2 demonstrates the case of multiple SNPs, taking $m = 3$ as an example for simplicity. In addition to the schemes demonstrated in Fig. 1, Fig. 2 shows two more rules. First, it includes the extra edge weights from each H_0^{Di} node to all non-parent H_0^{Lj} nodes, i.e. to all H_0^{Lj} with $j \neq i$. This allows for additional α -sharing between SNPs when both hypotheses (i.e. H_0^{Li} and H_0^{Di}) are rejected for any given SNP i . Second, the α -level is split with a Bonferroni type allocation between the m top-level hypotheses while none of α is initially provided to the m lower-level hypotheses. Upon rejection of a higher-level hypothesis, the lower-level child hypothesis receives all of the α/m -level of the parent (edge weight of 1). If the lower-level child hypothesis is then also found to be significant, its α threshold is then shared between all remaining higher-level hypotheses (edge weights of $1/2$).

The power advantage of our proposed GBA over the Bonferroni method is evident from the larger thresholds.

Where the Bonferroni method would test each hypothesis at the $\alpha/6$ -level, the GBA tests each hypothesis by thresholds that are no smaller than $\alpha/3$. To demonstrate, assume that H_0^{L1} and H_0^{L3} from Fig. 2 are rejected at the $\alpha/3$ level, but that H_0^{L2} is not. Then nodes corresponding to the rejected hypotheses H_0^{L1} and H_0^{L3} are removed and all α thresholds and edge weights are updated as shown in Fig. 3. Notice in Fig. 3 the reconnecting of edge weights which previously attached to H_0^{L1} and H_0^{L3} from H_0^{D2} , H_0^{D1} , and H_0^{D3} . This demonstrates how edges determine not only the weight that will be passed, but also define the inheritance of edge weights.

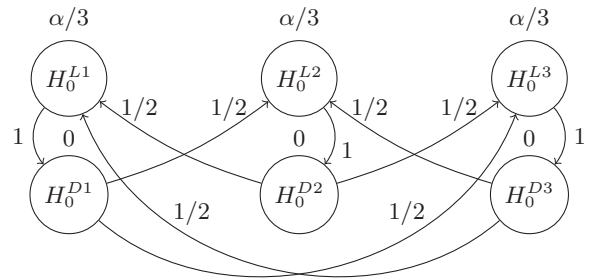


Fig. (2). Demonstration of the hierarchy of the GBA testing scheme for three SNPs.

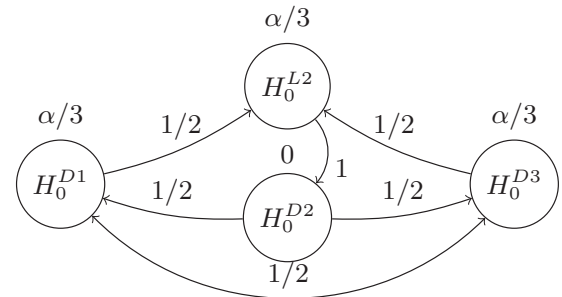


Fig. (3). Demonstration of the GBA testing scheme for three SNPs assuming that hypotheses H_0^{L1} and H_0^{L3} from the initial graph in (Fig. 2) are rejected.

Assume now that H_0^{D1} of Fig. 3 can be rejected at the $\alpha/3$ -level. The graph updating (Fig. 4A) becomes more complicated with this rejection because the rejected hypothesis is both sending out and taking in edge weight from the same hypotheses (nodes). Specifically, H_0^{D1} is set to send half of its threshold to H_0^{D3} and the other half to H_0^{D1} . Of the half that the now rejected H_0^{D1} would have received from H_0^{D2} , half is designated to H_0^{L2} and the other half designated to go to H_0^{D3} . This assignment causes the updated H_0^{D2} to send a total weight of $3/4$ to H_0^{D3} . However, recalling the logical structure of the hypotheses, it can be seen that H_0^{D2} will not be considered for testing unless H_0^{L2} is first rejected. Hence, the $1/4$ that H_0^{D2} would pass on to H_0^{L2} through H_0^{D3} at this point is not logically

possible as this would require testing H_0^{D2} before testing H_0^{L2} . This logical restriction allows us to move the 1/4 out from H_0^{D2} by means of the only other path available, so that H_0^{D3} receives a total weight of 1 from H_0^{D2} , as shown in Fig. 4A.

The node corresponding to H_0^{D3} in Fig. 3 was sending half of its threshold to H_0^{D1} and the other half to H_0^{L2} . With the removal of H_0^{D1} , now assumed to be significant, H_0^{D3} will now be doubly joined to H_0^{L2} and to itself by inheriting the outgoing paths from H_0^{D1} to both H_0^{L2} and H_0^{D3} . This junction of H_0^{D3} to itself would specify that the 1/2 that was going from H_0^{D3} to H_0^{D1} times the 1/2 that was going from H_0^{D1} to H_0^{D3} would result in H_0^{D3} returning 1/4 to itself. Since it is not possible for H_0^{D3} to pass 1/4 back to itself, it passes to H_0^{L2} the original 1/2 it was already sending to H_0^{L2} , plus the 1/4 inherited by H_0^{L2} from H_0^{D3} via H_0^{D1} plus the 1/4 that H_0^{D3} re-inherited from H_0^{D1} . The result is to have H_0^{D3} send all of its threshold to H_0^{L2} . This can also be viewed more simply by the fact that upon removal of H_0^{D1} from the graph (due to its rejection), H_0^{D3} is left with only one outgoing edge to H_0^{L2} , hence all of its threshold must be passed to H_0^{L2} .

The final graph resulting from the rejection of H_0^{D1} in Fig. 3 is depicted in Fig. 4A. At this point it could be possible that H_0^{L2} is rejected, but to demonstrate a more interesting scenario, assume that H_0^{D3} only can be rejected at the $\alpha/2$ -level. The resulting graph with H_0^{D3} removed is depicted in Fig. 4B. Interestingly, both H_0^{L2} and (if significant) H_0^{D2} can now be tested at the full level α .

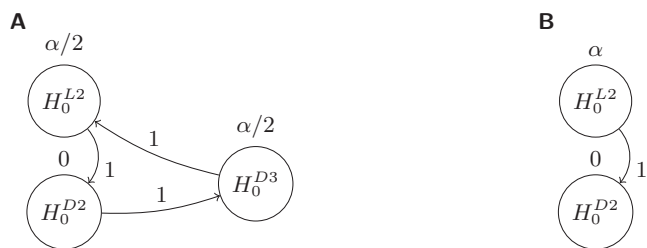


Fig. (4). **A:** The updated graph from (Fig. 3) assuming the hypothesis H_0^{D1} of (Fig. 3) is rejected at the $\alpha/3$ -level. **B:** Graph resulting from the rejection of the hypothesis H_0^{D3} at the $\alpha/2$ -level.

In conclusion, we solve the three problems in the multiplicity testing scenario that exist in previous LD based

QTL models. Firstly, the unidentifiable issue of D under the null hypothesis $H_0^L : \mu_1 = \mu_2 = \mu_3$ is avoided by skipping testing H_0^D when H_0^L is not rejected. Secondly, by flexibly passing out different portions of significance level α according to the conclusion of other tests, we make a correction for all multiple SNPs. Thirdly, by setting H_0^L to be the primary test over H_0^D , we satisfy the real application concern that existence of QTL is the ultimate goal. In the remainder of this article, we show through simulation studies that the proposed GBA is more powerful for LD based QTL mapping than standard Bonferroni adjustments and thus leads to more scientific discovery while maintaining strong control of the FWER.

RESULTS

Power Simulation

We investigated a simulation study to quantify the power advantage of the proposed graphical Bonferroni approach (GBA) over the standard Bonferroni adjustment within the LD based QTL mapping model [13]. The QTL, phenotype, and SNPs were generated under the assumptions of the alternative hypotheses (described in Methodology section in (2) and (3)). The QTL was generated using an assigned probability of $q = 0.7$ for the major allele. For each individual i , $Q_i = l$ with $l \in \{1,2,3\}$ was used to code the QTL genotypes of aa , Aa , and AA , respectively. The normally distributed phenotype dependent on the value of the QTL is generated as $Y_i | (Q_i = l) \sim N(\mu_l, \sigma)$, $i = 1, \dots, n$. The means for the phenotype Y corresponding to the values of the QTL were set at $\mu_1 = 8$, $\mu_2 = 10$ and $\mu_3 = 12$. SNPs were then generated using the conditional probability of the SNP genotype given the value of the QTL genotype for each individual. In general, for an LD based QTL mapping model, researchers genotype the SNP first and then use the SNP to generate a QTL based on the conditional probability of QTL genotype given SNP genotype. However, for our purposes, we are interested in extending from single SNP mapping to multiple SNPs mapping. Therefore, we derive the conditional probability of SNP genotype given QTL genotype (see Table 1) from the Bayes Rule in Equation (6).

$$P(M | QTL) = \frac{P(QTL | M)P(M)}{P(QTL)} \tag{6}$$

Sample sizes of $n = 100, 300$, and 500 were used to represent small, medium, and large sample sizes, respectively. The number of SNPs per simulation was set at $m = 1, 10, 50, 100, 500$, and $1,000$ to show the initial power under the single SNP scenario and the corresponding decreasing power trend as the number of SNPs increases. Finally, the heritability was set at two values, $H^2 = 0.1$ and 0.4 , corresponding to high and low error variance [27]. The model error variance σ^2 was computed using the heritability and genetic variance of the QTL. Power estimates were averaged over 1,000 simulations.

Table 1. The theoretical conditional probabilities of SNP genotype (columns) given QTL genotype (rows).

	<i>MM</i>	<i>Mm</i>	<i>mm</i>
<i>AA</i>	$\frac{p_{11}^2}{q^2}$	$\frac{2p_{11}p_{01}}{q^2}$	$\frac{p_{01}^2}{q^2}$
<i>Aa</i>	$\frac{2p_{11}p_{10}}{q(1-q)}$	$\frac{2(p_{11}p_{00} + p_{10}p_{01})}{q(1-q)}$	$\frac{2p_{10}p_{00}}{q(1-q)}$
<i>aa</i>	$\frac{p_{10}^2}{(1-q)^2}$	$\frac{2p_{10}p_{00}}{(1-q)^2}$	$\frac{p_{00}^2}{(1-q)^2}$

The simulation results, shown in Table 2 and depicted in Fig. 5, demonstrate the power comparison of the proposed GBA with the traditional Bonferroni adjustment. The traditional Bonferroni adjustment approach makes corrections not only for multiple SNPs but also for two tests of each SNP, with total $2m$ number of tests. Here m is number of SNPs. Our proposed GBA performs number of corrections somewhere between m and $2m$, flexibly depending on the real situation of each SNP. The designed simulation results also provides an experimental reference for researchers about how power varies among different sample size n , the number of SNPs m , and the degree of heritability (H^2). As expected, the power under high

heritability (B: $H^2 = 0.4$) is much higher than that of the low heritability (A: $H^2 = 0.1$) and the power under large sample size ($n = 500$, blue curves) is much higher than that of the small sample size ($n = 100$, green curves). Under high heritability ($H^2 = 0.4$) and a larger sample size ($n = 500$), the power of the multiplicity adjustment remains high even as the number of SNPs becomes large ($m = 1,000$). However, in practice it is often expensive to collect so many sample measurements. It is worth mentioning that the power obtained from the GBA can achieve 80% for a large number of SNPs ($m = 1,000$) but medium sample size ($n = 300$). Moreover, even with a low heritability ($H^2 = 0.1$), the power increase of the GBA over the Bonferroni adjustment allows for the possibility of maintaining the power level of the Bonferroni adjustment while decreasing the sample size of the study or increasing the number of SNPs, a great advantage for researchers. For example, under many cases, the power of our improved approach for 1,000 SNPs is similar or larger than the power of the Bonferroni adjustment for 500 SNPs.

Although the power increase of our proposed method improves moderately over the standard Bonferroni adjustment for the case of low heritability ($H^2 = 0.1$) when the sample size is small ($n = 100$), the power gains are still comparable to seminal results found by previous multiplicity improvements over their competitors [35]. All in all, our proposed GBA method shows a substantial increase in power over the Bonferroni adjustment under all 12 circumstances

Table 2. The results of the power simulation as depicted in Figure 5.

m	n = 100		n = 300		n = 500	
	Bon.	GBA	Bon.	GBA	Bon.	GBA
$H^2 = 0.1$						
1	0.333	0.422	0.604	0.696	0.753	0.831
10	0.132	0.207	0.313	0.462	0.490	0.677
50	0.062	0.093	0.186	0.272	0.324	0.465
100	0.045	0.067	0.146	0.210	0.265	0.379
500	0.022	0.032	0.079	0.112	0.163	0.229
1000	0.016	0.023	0.060	0.085	0.130	0.180
$H^2 = 0.4$						
1	0.751	0.825	0.984	0.992	1.000	1.000
10	0.500	0.671	0.929	0.990	0.994	1.000
50	0.340	0.480	0.853	0.975	0.983	1.000
100	0.283	0.396	0.811	0.955	0.974	0.999
500	0.180	0.248	0.699	0.862	0.944	0.998
1000	0.145	0.198	0.647	0.806	0.925	0.994

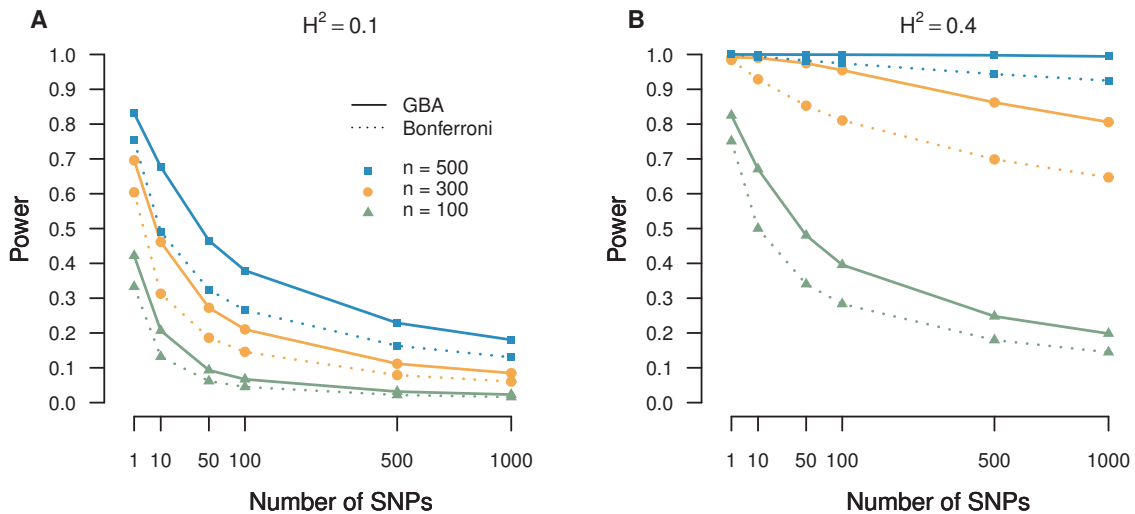


Fig. (5). Power comparison between the graphical Bonferroni approach (GBA) and standard Bonferroni adjustment under different sample size, number of SNPs, and heritability (**A:** $H^2 = 0.1$, **B:** $H^2 = 0.4$).

with the different combinations of sample size, number of SNPs, and heritability. We did not compare with permutation because it is computationally infeasible to compute these six settings using permutation. In addition, permutation approach is designed for the one testing structure, which is not the case here. Finally, we firmly believe that the GBA works better in the LD based QTL model because of the priority order setting, the logical consistency, uniformly better power in theory, and increase in computational efficiency.

Mouse HDL Cholesterol QTL Mapping Project

Epidemiological studies have consistently shown that the level of plasma high density lipoprotein (HDL) cholesterol is negatively correlated with the risks of coronary artery disease and gallstones [36, 41]. Because of the inverse relationship between HDL and cardiovascular disease, there has been considerable interest in understanding genetic mechanisms contributing to variations in HDL levels. HDL levels vary considerably in different people, which are affected by interactions of multiple genes and environmental factors, and up to 70% of this variation in humans is genetically determined [37, 42]. Because of the concordance between human QTLs regulating HDL and corresponding mouse loci and many easily controlled experimental advantages, mouse has become an animal model in HDL research. Numerous findings in HDL QTL associations are obtained from crosses between different inbred mouse strains. By crossing inbred strains that significantly differ in HDL levels and subsequently testing for association between HDL levels and genetic SNPs in the progeny, numerous significant QTLs involved in HDL have been identified in mouse [36, 41, 43, 47].

Compared to the inbred mice strains with coarse mapping resolution, the QTL research on wild-caught and commercial stocks of outbred mice, as resources for genetic fine mapping, is far under developed. Zhang *et al.* published an open resource outbred mouse database with 288 Naval Medical Research Institute (NMRI) mice and 44,428 unique SNP genotypes (available at <http://cgd.jax.org/datasets/>

[datasets.shtml](#)) [48]. Three hundred 4-to-6-week-old male NMRI mice were purchased and individually housed with the same diet and environmental conditions. The blood samples of each mouse were measured by submandibular puncture after a 4-hr fast. Then plasma samples were frozen for measurement of HDL cholesterol. There were 10 mice removed because the standard deviation of individual blood pressure is greater than two. Another two mice were also discarded for their 99% identity of SNP genotypes. This caused the final sample size to be 288. A total of 581,672 high density SNP were initially genotyped by the Novartis Genomics Factory using the Mouse Diversity Genotyping Array [49]. In order to guarantee promising data for association mapping studies [50], only polymorphic SNPs with minor allele frequency greater than 2%, Hardy-Weinberg equilibrium $\chi^2 < 20$, and missing values less than 40% were retained. Moreover, identical SNPs within a 2Mb interval were collapsed. This left 44,428 unique SNP genotypes for their resulting analysis using three analysis methods, linear trend test, two way ANOVA, and EMMA [51]. From Zhang's work, adjustments for multiplicity at the genome-wide association level were made using a simulation approach [52] as well as the permutation approach [53]. They identified three loci as significant, with two loci on Chromosome1 (Chr1) and a single locus on Chromosome5 (Chr5) (see Fig. 3 of [48]). However, after a closer investigation, Zhang *et al.* reported that the significant findings in Mb182 of Chr1 is spurious.

We applied the introduced LD based QTL model [13] and the proposed GBA multiplicity correction approach to this outbred mouse HDL cholesterol genome data to compare our findings with the highly validated discoveries in current literature. Recalling the detailed adjustment structure of the GBA, it can be seen that the adjusted p -value obtained from GBA for the test of H_0^D will never be smaller than that of H_0^L . Hence, reporting the significant adjusted p -values for H_0^D is sufficient for demonstrating those SNPs that show

strong evidence of linkage to a significantly existed QTL. (Fig. 6) depicts the negative log of the adjusted p -values for H_0^D for each SNP as a function of the location (in Mb) of each SNP for 19 chromosomes. The threshold for the adjusted p -values of $-\log(0.05) \approx 2.9957$ supports two dramatically significant findings, on Chr1 at Mb173 and Chr5 at Mb125. These two significant discoveries are the same as the findings in current outbred mouse literature, compare to Fig. 3 of [48], but with an even stronger signal.

In Table 3 we notice that all significant QTLs detected from outbred mouse by our model are confirmed from reported findings obtained from inbred mouse crosses using very different approaches. Two QTLs have been reported coincident with candidate genes. Chr1 locus at Mb173, the highest peak in Fig. 6, is the major determinant of HDL, which has been detected as QTL *Hdlq15* in inbred mouse strains multiple times (as referenced in Table 3). Combining mouse crosses with haplotype analysis for the HDL QTL located on Chr 1 locus at Mb173 reduced the list of candidates to a small amount. Numerous mouse crosses have linked HDL to this region, and *Apoa2* has been identified as the gene underlying the QTL [38, 40, 41, 43, 46]; this gene has been highlighted in *Nature Reviews Genetics* [54]. Chr5 locus at Mb125, the second highest peak in Fig. 6, is located in the same locus as QTL *Hdlq1* found by [46] and [44] (as referenced in Table 3). In addition, they conclude that *Scarb1* (a well known gene involved in HDL metabolism) is the causal gene underlying *Hdlq1* by haplotype analysis, gene sequencing, expression studies, and a spontaneous mutation [45, 47].

There are a total of two significant detections from a large pool of 44,428 candidate SNPs from our model, with the findings confirmed by inbred mouse findings. Zhang *et*

al. worked on exactly the same data, adjusting for multiplicity at the genome-wide association level using a simulation approach [52] as well as the permutation approach [53]. They made the same two significant detections with less significance level than our p -values, and they also find one spurious QTL. Therefore, our proposed approach brings a useful alternative approach for SNP selection literature.

DISCUSSION

Detecting significant genes that cause disease (for example the inverse relation between human cholesterol and cardiovascular disease) or regulate biological traits through LD based QTL mapping has been popular in many disciplines [1-13]. The new techniques can simultaneously consider tens of thousands of SNPs and hence bring big challenges to multiple testing. In addition, high dimensional biological traits, often reduced to multiple PC components, have been widely used and add yet another demand for a powerful and computationally efficient approach to adjust for multiple tests [13, 55-57].

These multiple tests require an adjustment on the resulting p -values in order to preserve control of the familywise error rate (FWER) at a pre-specified level α . In some cases, follow up work on the significant findings may justify using the false discovery rate (FDR) as the error rate of interest. Typically however, the significant results are directly reported and therefore the FWER is the more desirable form of error rate to control. The current standard approach in LD based QTL mapping is to apply a Bonferroni adjustment to correct for multiplicity and preserve the FWER. As is well known, the Bonferroni correction is overly conservative for large numbers of tests, but the

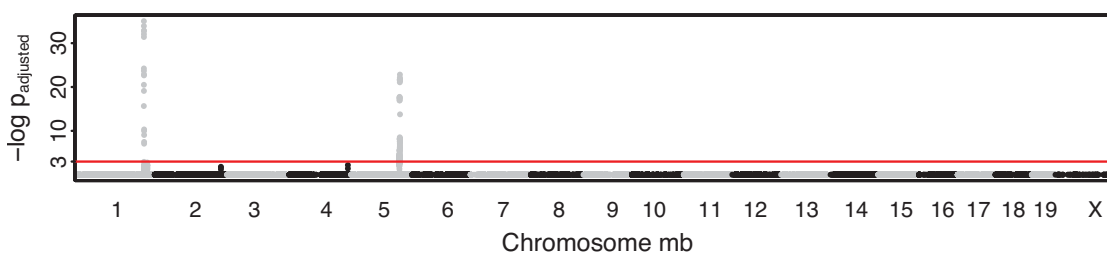


Fig. (6). The negative log of the GBA-adjusted p -values for H_0^D for each SNP in the mouse HDL cholesterol QTL mapping project. The red reference line corresponds to a 0.05 familywise error rate.

Table 3. The significant results of the outbred mice HDL cholesterol QTL mapping project depicted in Figure 6. SNPs are ordered by significance level. Corresponding concurrence candidate gene and QTL from previous inbred crosses studies are shown.

Chr	Position (Mb) 5pt	Adjusted P 5pt	Raw P 5pt	Raw P	Candidate	Inbred	Ref.
		(H_0^D)	(H_0^L)	(H_0^D)	Gene	QTL	
1***	173,155,512	5.7×10^{-15}	1.3×10^{-19}	3.0×10^{-30}	<i>Apoa2</i>	<i>Hdlq15</i>	[38,41,46,47,54]
5***	125,530,593	5.2×10^{-10}	1.2×10^{-14}	2.0×10^{-83}	<i>Scarb1</i>	<i>Hdlq1</i>	[37,40,44-47]

*** Significant at the FWER $5 * 10^{-10}$ level.

advantages of simplicity without independence assumptions on the corresponding family of tests continue to make it popular.

In this article, we tailored a multiple correction approach, based on graphical weighted-Bonferroni methods [25], which allows for the logical order among the two hypotheses to be structured into the multiplicity correction. As in the LD based QTL mapping model of [13], we need to test two hypotheses for each SNP, one with H_0^L about whether or not an association exists between QTL and phenotype, and the other with H_0^D about whether or not LD exists between SNP and QTL. Among these two tests, the QTL existence test has higher priority because the LD test will not be applicable if a QTL does not exist, and the existence of QTL is the ultimate goal in real applications. Although the logical structure of the two tests is known, none of the current LD based QTL literature considers this priority structure when performing these two tests [13, 23, 24, 27]. By the structure of graphical weighted Bonferroni, the quantitative trait loci test and linkage disequilibrium test are integrated into a combined multiple testing correction framework [58]. As a result, GBA approach has more potential applications in QTL studies. For example, in a haplotype study, we can put QTL test, dominant or additive effect test into one multiple testing correction framework. Hence, if QTL test is not significant, we don't have to test dominant or additive effects.

The significance of the power advantage of the proposed method, established through simulations, and finally on real data, is such that we advocate its use whenever multiple tests are needed for the LD based QTL mapping design, where both H_0^L and H_0^D tests are considered.

CONFLICT OF INTEREST

The author(s) confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

This work is supported by corresponding author's Utah State University VPR Research Catalyst Grant.

REFERENCES

[1] Terwilliger, J.D.; Weiss, K.M. Linkage disequilibrium mapping of complex disease: Fantasy or reality? *Curr Opin Biotechnol*, **1998**, *9*, 578-594.

[2] Service, S.K.; Lang, T.D.W.; Freimer, N.B.; Sandkuijl, L.A. Linkage-disequilibrium mapping of disease genes by reconstruction of ancestral haplotypes in founder populations. *Am. J. Hum. Genet.*, **1998**, *64*, 1728-1738.

[3] Martin, E.R.; Lai, E.H.; Gilbert, J.R.; Rogala, A.R.; Afshari, A.J. SNPing away at complex diseases: Analysis of single-nucleotide polymorphisms around APOE in Alzheimer disease. *Am. J. Hum. Genet.*, **2000**, *67*, 383-394.

[4] Goldstein, D. Weale M. Linkage disequilibrium holds the key. *Curr. Biol.*, *11*, 576-579.

[5] Sachidanandam, R.; Weissman, D.; Schmidt, S.C.; Kakol, J.M.; Stein, L.D. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, **2001**, *409*, 928-933.

[6] Reich, D.; Cargill, M.; Bolk, S.; Ireland, J.; Sabeti, P.C. Linkage

disequilibrium in the human genome. *Nature*, **2001**, *411*, 199-204.

[7] Ardlie, K.; Kruglyak, L.; Seielstad, M. Patterns of linkage disequilibrium in the human genome. *Nat. Rev. Genet.*, **2002**, *3*, 299-309.

[8] Dawson, E.; Abecasis, G.R.; Bumpstead, S.; Chen, Y.; Hunt, S. A first generation linkage disequilibrium map of human chromosome 22. *Nature*, **2002**, *418*, 544-548.

[9] Morris, R.W.; Kaplan, N.L. On the advantage of haplotype analysis in the presence of multiple disease susceptibility alleles. *Genet Epidemiol.*, **2002**, *23*, 221-233.

[10] Weiss, K.M.; Clark, A.G. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet.*, **2002**, *18*, 19-24.

[11] Skipper, L.; Wilkes, K.; Toft, M.; Baker, M.; Lincoln, S. Linkage disequilibrium and association of MAPT H1 in Parkinson disease. *Am. J. Hum. Genet.*, *2004* *75*, 669-677.

[12] Wang, L.; Luzynski, K.; Pool, J.E.; Janousek, V.; Dufkova, P. Measures of linkage disequilibrium among neighbouring SNPs indicate asymmetries across the house mouse hybrid zone. *Mol. Ecol.*, **2001**, *20*, 2885-3000.

[13] Fu, G.; Bo, W.; Pang, X.; Wang, Z.; Chen, L. Mapping shape QTLs using a radius-centroid-contour model. *Heredity*, **2013**, *110*, 511-519.

[14] Mooney, S. Bioinformatics approaches and resources for single nucleotide polymorphism functional analysis. *Brief Bioinform.*, **2005**, *6*, 44-56.

[15] Doerge, R.W. Mapping and analysis of quantitative trait loci in experimental populations. *Nat. Rev. Genet.*, **2002**, *3*, 43-52.

[16] Li, J.; Das, K.; Fu, G.; Li, R.; Wu, R. The Bayesian lasso for genome-wide association studies. *Bioinformatics*, **2011**, *27*, 516-523.

[17] Nyholt, D.R. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am. J. Hum. Genet.*, **2004**, *74*, 765-769.

[18] Johnson, R.; Nelson, G.W.; Troyer, J.L.; Lautenberger, J.A.; Kessing, B.D. Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics*, **2010**, *11*, 724.

[19] Nakagawa, S. A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behav. Ecol.*, **2004**, *15*, 1044-1045.

[20] Gao, X.; Starmer, J.; Martin, E.R. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol.*, **2008**, *32*, 361-369.

[21] Han, B.; Kang, H.M.; Eskin, E. Rapid and accurate multiple testing correction and power estimation for millions of correlated SNPs. *PLoS Genet*, **2009**, *5*(4), e1000456 doi:10.1371/journal.pgen.1000456.

[22] Cheng, R.; Lim, J.E.; Samocha, K.E.; Sokoloff, G.; Abney, M. Genome-wide association studies and the problem of relatedness among advanced intercross lines and other highly recombinant populations. *Genetics*, **2010**, *185*, 1033-1044.

[23] Lou, X.; Casella, G.; Littell, R.C.; Yang, M.C.K.; Johnson, J.A. A haplotype-based algorithm for multilocus linkage disequilibrium mapping of quantitative trait loci with epistasis. *Genetics*, **2003**, *4*, 1533-1548.

[24] Das, K.; Wu, R. A statistical model for the identification of genes governing the incidence of cancer with age. *Theor. Biol. Med. Mod.*, **2008**, *5*, 7.

[25] Bretz, F.; Maurer, W.; Brannath, W.; Posch, M. A graphical approach to sequentially rejective multiple test procedures. *Stat. Med.*, **2009**, *28*, 586-604.

[26] Wu, R.; Ma, C.; Casella, G. Statistical Genetics of Quantitative Traits: Linkage, Maps and QTL. Springer-Verlag, New York, 1st edition. **2007**.

[27] Wang, Z.; Wu, R. A statistical model for high-resolution mapping of quantitative trait loci determining HIV dynamics. *Statist Med.*, **2004**, *23*, 3033-3051.

[28] Fu, G.; Berg, A.; Das, K.; Li, J.; Li, R. A statistical model for mapping morphological shape. *Theor. Biol. Med. Model*, **2010**, *7*, 28.

[29] Brown, A.H.D. Sample size required to detect linkage disequilibrium between two or three loci. *Theor. Popul. Biol.*, **1975**, *8*, 184-201.

[30] Hedrick, P.W. Gametic disequilibrium measures: Proceed with caution. *Genetics*, **1987**, *2*, 331-341.

[31] Pritchard, J.K. Przeworski, M. Linkage disequilibrium in humans: Models and data. *Am. J. Hum. Genet.*, **2001**, *69*, 1-14.

[32] Mueller, J.C. Linkage disequilibrium for different scales and applications. *Brief Bioinform.*, **2004**, *5*, 355-364.

[33] Davies, R.B. Hypothesis testing when a nuisance parameter is

- present only under the alternative. *Biometrika*, **1977**, *64*, 247-254.
- [34] Davies, R.B. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, **1987**, *74*, 33-43.
- [35] Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.*, **1995**, *Series B 57*, 298-300.
- [36] Mehrabian, M.; Castellani, L.; Wen, P.; Wong, J.; Rithaporn, T. Genetic control of HDL levels and composition in an interspecific mouse cross (CAST/Ei * C57BL/6J). *J. Lipid. Res.*, **2000**, *41*, 1936-1946.
- [37] Wang, X.; Le Roy, I.; Nicodeme, E.; Li, R.; Wagner, R. Using advanced intercross lines for high-resolution mapping of HDL cholesterol quantitative trait loci. *Genome Res.*, **2003**, *13*, 1654-1664.
- [38] Wang, R.K.D.H.X.; Paigen, B. Haplotype analysis in multiple crosses to identify a QTL gene. *Genome Res.*, **2004**, *14*, 1767-1772.
- [39] Lyons, M.; Wittenburg, H.; Li, R.; Walsh, K.; Churchill, G.A. Quantitative trait loci that determine lipoprotein cholesterol levels in DBA/2J and CAST/Ei inbred mice. *J. Lipid. Res.*, **2003**, *44*, 953-967.
- [40] Su, Z.; Ishimori, N.; Chen, Y.; Leiter, E.H.; Churchill, G.A. Four additional mouse crosses improve the lipid QTL landscape and identify Lipg as a QTL gene. *J. Lipid. Res.*, **2009**, *50*, 2083-2094.
- [41] Su, Z.; Wang, X.; Tsaih, S.W.; Zhang, A.; Cox, A. Farp2 and Stk25 are candidate genes for the HDL cholesterol locus on mouse chromosome 1. *Arterioscl. Throm. Vas.*, **2009**, *29*, 107-113.
- [42] Rader, D.J.; Maugeais, C. Genes influencing HDL metabolism: New perspectives and implications for atherosclerosis prevention. *Mol. Med. Today*, **2000**, *6*, 170-175.
- [43] Machleder, D.; Ivandic, B.; Welch, C.; Castellani, L.; Reue, K. Complex genetic control of HDL levels in mice in response to an atherogenic diet. *J. Clin. Invest.*, **1997**, *99*, 1406-1419.
- [44] Korstanje, R.; Li, R.; Howard, T.; Kelmenson, P.; Marshall, J. Influence of sex and diet on quantitative trait loci for HDL cholesterol levels in an SM/J by NZB/BINJ intercross population. *J. Lipid. Res.*, **2004**, *45*, 881-888.
- [45] Wergedal, J.E.; Ackert-Bicknell, C.L.; Beamer, W.G.; Mohan, S.; Baylink, D.J. Mapping genetic loci that regulate lipid levels in a NZB/B1NJ*RF/J intercross and a combined intercross involving NZB/B1NJ, RF/J, MRL/MpJ, and SJL/J mouse strains. *J. Lipid. Res.*, **2007**, *48*, 1724-1734.
- [46] Su, Z.; Wang, X.; Tsaih, S.W.; Zhang, A.; Cox, A. Genetic basis of HDL variation in 129/SvImJ and C57BL/6J mice: Importance of testing candidate genes in targeted mutant mice. *J. Lipid. Res.*, **2009**, *50*, 116-125.
- [47] Su, Z.; Leduc, M.S.; Korstanje, R.; Paigen, B. Untangling HDL quantitative trait loci on mouse chromosome 5 and identifying Scarb1 and Acads as the underlying genes. *J. Lipid. Res.*, **2010**, *51*, 2706-2713.
- [48] Zhang, W.; Korstanje, R.; Thaisz, J.; Staedtler, F.; Hartman, N. Genome-wide association mapping of quantitative traits in outbred mice. *G3 (Bethesda)*, **2012**, *14*: 167-174.
- [49] Yang, H.; Ding, Y.; Hutchins, L.N.; Szatkiewicz, J.; Bell, T.A. A customized and versatile high-density genotyping array for the mouse. *Nat. Methods*, **2009**, *6(9)*, 663-666.
- [50] Yalcin, B.; Nicod, J.; Bhomra, A.; Davidson, S.; Cleak, J. Commercially available outbred mice for genome-wide association studies. *Plos Genetics*, **2010**, *6(9)*, e1001085.
- [51] Kang, H.M.; Zaitlen, N.A.; Wade, C.M.; Kirby, A.; Heckerman, D. Efficient control of population structure in model organism association mapping. *Genetics*, **2008**, *178*, 1709-1723.
- [52] Knijnenburg, T.A.; Wessels, L.F.; Reinders, M.; Shmulevich, I. Fewer permutations, more accurate p-values. *Bioinformatics*, **2009**, *25(12)*, 161-168.
- [53] Churchill, G.A.; Doerge, R.W. Empirical threshold values for quantitative trait mapping. *Genetics*, **1994**, *138*, 963-971.
- [54] Flint, J.; Eskin, E. Genome-wide association studies in mice. *Nat. Rev. Genet.*, **2012**, *13*, 807-817.
- [55] Langlade, N.; Feng, X.; Drans_eld, T.; Copey, L.; Hanna, A. Evolution through genetically controlled allometry space. *P. Natl. Acad. Sci. USA*, **2005**, *102*, 10221-10226.
- [56] Drake, A.; Klingenberg, C. Large-scale diversification of skull shape in domestic dogs: Disparity and modularity. *Am. Nat.*, **2010**, *175*, 289-301.
- [57] Fu, G.; Saunders, G.; Stevens, J.R. Holm multiple correction for large-scale gene-shape association mapping. *BMC Genetics*, **2014**, *15(Suppl 1)*, S5.
- [58] Saunders, G. Family-wise error rate control in QTL mapping and gene ontology graphs with remarks on family selection. Ph.D. thesis, Utah State University, Department of Mathematics and Statistics. **2014** <http://digitalcommons.usu.edu/etd/2164/>