A Unified Framework for Detecting Rare Variant Quantitative Trait Associations in Pedigree and Unrelated Individuals via Sequence Data

Dajiang J. Liu^{a, b} Suzanne M. Leal^{a, b}

^aDepartment of Molecular and Human Genetics, Baylor College of Medicine, and ^bDepartment of Statistics, Rice University, Houston, Tex., USA

Key Words

Extreme sampling • Next-generation sequencing • Pedigree samples • Quantitative trait loci • Rare variants

Abstract

Objectives: There is great interest to sequence unrelated or pedigree samples for detecting rare variant quantitative trait associations. In order to reduce the cost of sequencing and improve power, many studies sequence selected samples with extreme traits. Existing methods for detecting rare variant associations were developed for unrelated samples. Methods are needed to analyze (selected or randomly ascertained) pedigree samples. **Methods:** We propose a unified framework of modeling extreme trait genetic associations (MEGA) with rare variants. Using MEGA and appropriate permutation algorithms, many rare variant tests can be extended to family data. As an application, we compared study designs using both sib-pairs and unrelated individuals. Extensive simulations were carried out using realistic population genetic and complex trait models. **Results:** It is demonstrated that when extreme sampling is implemented within equal-sized cohorts of unrelated individuals or sib-pairs, analyzing unrelated individuals is consistently more powerful than studying sib-pairs. A higher portion of rare variants can

be identified through sequencing unrelated samples compared to sibs. Alternatively, if samples are ascertained using fixed thresholds from an infinite-sized population, sequencing one sib with the most extreme trait from each extreme concordant sib-pair is consistently the most powerful design. **Conclusions:** MEGA will play an important role in the analysis of sequence-based genetic association studies.

Copyright © 2012 S. Karger AG, Basel

Introduction

For the genetic etiology of complex traits (CT), there are 2 parallel hypotheses: CT are influenced by common variants (CV) that have modest phenotypic effect (CT/CV), or they can be due to rare variants (CT/RV) that have a larger phenotypic effect than CV [1]. There is solid evidence supporting the CT/RV hypothesis [2–7]. In order to map quantitative trait (QT) loci (QTL), linkage and association approaches have been widely applied. Linkage analysis methods can usually only map traits to large intervals consisting of many megabases of DNA. On the other hand, association analysis can be used to localize traits to more refined regions [8]. If the CT/RV hypothesis holds true, traditional indirect association mapping

will be underpowered by design, due to the low correlations (r^2) between common tagSNPs and rare QTL variants. Direct association mapping through sequencing and analyzing exomes or genomes is the optimal way to detect associations with RV [9]. With the development and application of next-generation sequencing technologies such as Illumina HiSeq, ABI SOLiD and Roche 454, direct association mapping of CT in large scale genetic studies has been made possible. However, it is still expensive to perform either exome or genome sequencing on a large number of individuals.

A large number of well-phenotyped cohorts of both related and unrelated individuals are available for genetic studies, which include the Mid-Atlantic Twin Registry [10], Dallas Heart Study [11] and Framingham Heart Study [12]. There is currently great interest to sequence samples from these cohorts, in order to better understand CT etiologies. Given the budgetary constraints, most studies can only afford to sequence a limited number of samples. Therefore an extreme QT study design is often implemented, in order to maximize power for a fixed sequence sample size. It was shown previously that using families with multiple affected individuals can be more advantageous for enriching causal variant alleles. However, it is not always more powerful to use related samples for detecting QT associations [13, 14]. The vast majority of QTL association studies are performed using existing cohorts of randomly ascertained individuals. For a fixed sample size, the power is strongly influenced by the size of the cohort and the selective sampling study designs. Using a smaller cohort or applying a more stringent selection criteria (e.g. the study design that requires both sibs in a sib-pair have extreme concordant QT value, QTV) will result in selecting individuals without particularly extreme QTV and reduce power. In practice, it is of interest to know whether it is more advantageous to use unrelated or pedigree samples when studies are implemented from existing cohorts of fixed size.

Many methods have been developed for mapping RV, such as the combined multivariate and collapsing (CMC) method [9], the test of the aggregated number of RV (ANRV) [15], the weighted-sum statistics (WSS) [16], the variable threshold tests (VT) [17], the RARECOVER method [18], the kernel-based adaptive cluster (KBAC) method [19], the replication-based test (RBT) [20], the c -alpha test [21], the sequence kernel association test (SKAT) [22], etc. However, these methods were exclusively developed for the analysis of unrelated individuals.

A couple of methods [23, 24] use information from external samples of sib-pairs to determine weights or high-

risk haplotypes and then test for RV associations in an independent sample of unrelated individuals. Specifically, the method of Zhu et al. [25] uses affected sib-pairs to screen and identify high-risk haplotypes, while the method by Feng et al. [26] utilizes affected or discordant sib-pairs to assign weights to different variant sites. Both methods do not model phenotypic or genetic correlations between related individuals in association testing and they cannot be used to analyze pedigree samples.

In this article, a unifying mixed effects likelihood framework modeling extreme trait genetic associations (MEGA) is proposed for direct QTL mapping using unrelated individuals and pedigree samples with extreme traits. MEGA modifies the commonly used variance component model, and the QTL effects are modeled as fixed effect to facilitate joint analysis of multiple RV [23, 24]. In order to accommodate complicated ascertainment mechanisms, a prospective likelihood method is coupled with the mixed effects model. MEGA can be applied to any extreme QT study of unrelated individuals and/or pedigrees. Association testing can be performed using score statistics, which is asymptotically normally distributed when regularity conditions are satisfied. When normality fails to hold, appropriate permutation algorithms were developed, which can be applied to obtain p values empirically for simple pedigrees such as nuclear families. Combined with the permutation algorithms, many RV association methods that require evaluating p values empirically can be extended to small pedigrees.

Applying MEGA, we investigated optimal study design strategies using unrelated individuals and sib-pairs. Previously for sib-pair study designs some comparisons were made for detecting CV extreme QT associations [8, 27, 28]. However, phenotypic and population-genetic models are very different for CV and RV associations. In addition, previous comparisons were focused on robust tests of associations between QT and the transmissions of alleles within families and did not consider studies using unrelated samples. With the development of advanced computational and statistical tools, such as EMMAX [24], population stratifications and cryptic relatedness can also be effectively controlled in population-based studies. It is of interest to compare the power for study designs using unrelated individuals or family data.

The power for detecting associations using unrelated individuals and sib-pairs was evaluated. To carry out selective sampling, 2 thresholds for QTVs are chosen, i.e. the upper bound y^{ub} and lower bound y^{lb} , and individuals with QTVs exceeding these thresholds are selected.

Study Designs Using Unrelated Individuals

- (1) Extreme unrelated individuals (EUI): unrelated individuals with QTVs $> y^{ub}$ or $< y^{lb}$ are selected.

Study Designs Using Sib-Pairs

Study Designs Where Phenotype and Sequence Data from Both Sibs Are Analyzed

- (2) Extreme Concordant Sib-Pair (ECSP): sib-pairs are selected with both sibs having 'concordant' extreme trait values, i.e. the QTVs for both sibs are either $> y^{ub}$ or $< y^{lb}$.
- (3) Extreme Discordant Sib-Pair (EDSP): this design was previously discussed by Risch and Zhang [29, 30], Carey and Williamson [31] and Eaves and Meyer [32]. Sib-pairs are selected with both sibs having 'discordant' extreme trait values, i.e. the QTV for one sib is $> y^{ub}$, while the QTV for the other sib is $< y^{lb}$.
- (4) Extreme Discordant and Concordant Sib-Pair (EDAC): this design was first described by Gu et al. [33]. Sib-pairs are selected with both sibs having 'concordant' or 'discordant' extreme trait values, i.e. the QTVs for both sibs are either $> y^{ub}$ or $< y^{lb}$, or one sib has a QTV $> y^{ub}$, while the other sib has a QTV $< y^{lb}$.
- (5) Extreme Proband (EP): this design was first described by Abecasis et al. [27]. Sib-pairs are selected if at least one sib in the sib-pair has an extreme trait value ($> y^{ub}$ or $< y^{lb}$). The sib in each selected sib-pair with the most extreme trait value is designated as the 'proband'.

Study Designs Where Only One Sib Is Sequenced but Phenotype Data from Both Sibs Is Analyzed

- (6) Extreme Sib from an Extreme Concordant Sib-Pair (ES-ECSP): the sib with the most extreme trait value is sequenced from each extreme concordant sib-pair.
- (7) Extreme Sib from a Sib-Pair Selected by Proband (ES-EP): for each sib-pair selected by proband, only the proband is sequenced.

The mathematical definitions for each of the 7 study designs can be found in table 1, and the designs are also graphically illustrated in figure 1.

Two major categories of extreme QT designs are distinguished. First, in most genetic studies, samples are selected from existing finite-sized cohorts of sib-pairs or unrelated individuals. This type of selection is named 'cohort selection'. For the second type of selection strategy ('population ascertainment'), new samples are ascertained and screened from a population of unlimited size. The cutoffs used to screen samples from the general population can be determined by quantities of clinical relevance, such as the blood pressure thresholds used to di-

Fig. 1. Graphical illustration of the study designs evaluated. Red (black in the printed version) squares represent individuals with high extreme QTV. Blue (white in the printed version) squares represent individuals with low extreme QTV. Green (grey in the printed version) squares correspond to phenotyped individuals with any QTV. The arrows indicate which individuals are sequenced for each study design.

agnose hypertension, or the body mass index cutoffs that are used to classify individuals as being overweight or obese.

The aforementioned study designs were evaluated using a rigorous population-genetic framework. Site frequency spectrums (SFS) for Africans were generated using forward time simulations [34]. QTs were simulated using parameters estimated from clinically relevant complex phenotypes.

Based upon our extensive simulations, it is shown that for selections implemented within equal-sized cohorts of unrelated individuals and sib-pairs, sequencing unrelated individuals can be more powerful than sequencing an equal number of sib-pairs. In particular, EUI (design 1) is more powerful than all sib-pair study designs. In practice, if the size of the available cohort of unrelated individuals is larger, the advantage of analyzing unrelated individuals can be even more pronounced. For sib-pair co-

Table 1. Definitions of QT value thresholds for the study designs using unrelated individuals and sib-pairs

Study designs	Design No. / acronym	QT value threshold
<i>Unrelated individuals</i>		
Extreme unrelated individual	1. EUI	$\Omega_{EUI} = \{Y Y \geq y_{EUI}^{ub}\} \cup \{Y Y \leq y_{EUI}^{lb}\}$
<i>Sib-pairs</i>		
Extreme concordant sib-pair	2. ECSP	$\Omega_{ECSP} = \{(Y_1, Y_2) Y_1 \geq y_{ECSP}^{ub}, Y_2 \geq y_{ECSP}^{ub}\} \cup \{(Y_1, Y_2) Y_1 \leq y_{ECSP}^{lb}, Y_2 \leq y_{ECSP}^{lb}\}$
Extreme discordant sib-pair	3. EDSP	$\Omega_{EDSP} = \{(Y_1, Y_2) Y_1 \geq y_{EDSP}^{ub}, Y_2 \leq y_{EDSP}^{lb}\} \cup \{(Y_1, Y_2) Y_1 \leq y_{EDSP}^{lb}, Y_2 \geq y_{EDSP}^{ub}\}$
Extreme discordant and concordant sib-pair	4. EDAC	$\Omega_{EDAC} = \{(Y_1, Y_2) Y_1 \geq y_{EDAC}^{ub}, Y_2 \leq y_{EDAC}^{lb}\} \cup \{(Y_1, Y_2) Y_1 \leq y_{EDAC}^{lb}, Y_2 \geq y_{EDAC}^{ub}\}$ $\cup \{(Y_1, Y_2) Y_1 \geq y_{EDAC}^{ub}, Y_2 \geq y_{EDAC}^{ub}\} \cup \{(Y_1, Y_2) Y_1 \leq y_{EDAC}^{lb}, Y_2 \leq y_{EDAC}^{lb}\}$
Extreme proband	5. EP	$\Omega_{EP} = \{(Y_1, Y_2) Y_1 \geq y_{EP}^{ub}\} \cup \{(Y_1, Y_2) Y_2 \geq y_{EP}^{ub}\}$ $\cup \{(Y_1, Y_2) Y_1 \leq y_{EP}^{lb}\} \cup \{(Y_1, Y_2) Y_2 \leq y_{EP}^{lb}\}$
Extreme sib per extreme concordant sib-pair	6. ES-ECSP	$\Omega_{ES-ECSP} = \{(Y_1, Y_2) Y_1 \geq Y_2 \geq y_{ES-ECSP}^{ub}\} \cup \{(Y_1, Y_2) Y_1 \leq Y_2 \leq y_{ES-ECSP}^{lb}\}$
Extreme sib per sib-pair selected by extreme proband	7. ES-EP	$\Omega_{ES-EP} = \{(Y_1, Y_2) Y_1 \geq y_{ES-EP}^{ub}, Y_1 \geq Y_2\} \cup \{(Y_1, Y_2) Y_1 \leq y_{ES-EP}^{lb}, Y_1 \leq Y_2\}$

horts, sequencing individuals that are unrelated is still desirable, i.e. using sib-pairs selected by proband and sequencing only the proband from each selected sib-pair (design 6: ES-EP) is the most powerful study. In addition, sequencing unrelated samples also facilitates the identification of a greater number of novel variants, because of the increase in the number of independent chromosomes.

When new samples are ascertained from the general population, screening and sequencing pedigree samples with multiple individuals having extreme traits can be substantially more powerful than studying unrelated individuals. Specifically, (1) ES-ECSP (design 6) is consistently more powerful than analyzing unrelated individuals with extreme trait values (design 1: EUI), but ES-ECSP requires screening a very large number of sib-pairs. (2) For QT with low residual correlations, ECSP (design 2), where sib-pairs with concordant extreme traits are sequenced, or EDAC (design 4), where sib-pairs with concordant and discordant extreme trait values are sequenced, are both more powerful than using unrelated individuals (EUI). However, ECSP and EDAC have com-

parable power with EUI when trait heritability is high (>90%).

In order to further illustrate applications of MEGA, the method was applied to the analyses of a sequence dataset generated by the Ottawa Obesity Study, in which 56 candidate genes for obesity were sequenced using DNA samples from 378 obese and 379 lean individuals [2]. The results confirm previous analyses [2], and strengthen our conclusions.

Materials and Methods

A unifying framework is introduced for analyzing QT in sequencing-based association studies using samples from unrelated individuals and pedigrees with extreme QT phenotypes. The framework is based upon the mixed model approach which generalizes the Fisher's biometrical model [35, 36]. The locus-genetic effect is modeled as fixed effect such that multiple RV in a gene locus can be jointly analyzed. The correlation structure between pedigree members in the mixed model also differs from the variance components model.

When applied to extreme QTs, the mixed effects model is coupled with a prospective ascertainment-corrected likelihood approach. The framework incorporates QT information from all available pedigree members that are involved in the sample ascertainment. It also allows for efficient inferences of genetic parameters of interest for any study design. Appropriate permutation algorithms were developed for nuclear pedigrees, e.g. sib-pairs, such that most RV association methods developed for unrelated individuals can be incorporated.

Mixed Effects Model

The multi-site gene locus genotype for an individual is coded as a vector $\vec{X} = (x^1, x^2, \dots, x^s)$. Each x^s is coded by an indicator of whether the individual carries one or two RV at site s . For example, it takes value 1 if the genotype at nucleotide site s is either homozygous for the rare allele or heterozygous.

For human CT, each genetic locus only contributes a very small fraction of the total genetic variance [23]. In particular, the locus-genetic variation that is attributable to RV can be even smaller due to their low aggregated frequencies and moderate effect sizes. Therefore, the locus-genetic effect can be modeled as a fixed effect. This is similar to the ideas presented by Chen and Abecasis [37] and Kang et al. [24].

Based upon Fisher's biometrical model [35, 36], the mixed effects QT model for an 'individual' j from pedigree i is given by

$$Y_{ij} = \beta_0 + \beta_1 C(\vec{X}_{ij}, Y_{ij}) + g_{ij} + e_{ij} \quad (1)$$

The QT residual variation is partitioned into a polygenic effect g_{ij} , and an environmental effect e_{ij} , i.e. $g_{ij} \sim N(0, \sigma_g^2)$, $e_{ij} \sim N(0, \sigma_e^2)$.

The covariance matrix Σ_i of QTVs for N_i individuals from pedigree i in the mixed effects model is different from Fisher's variance components model [36], in that no identical by descent coefficients are involved. The covariance matrix satisfies

$$\Sigma_i = \begin{pmatrix} \sigma_g^2 + \sigma_e^2 & \pi_{i,1,2}\sigma_g^2 & \cdots & \pi_{i,1,N_i}\sigma_g^2 \\ \pi_{i,2,1}\sigma_g^2 & \sigma_g^2 + \sigma_e^2 & \cdots & \pi_{i,2,N_i}\sigma_g^2 \\ \vdots & \cdots & \cdots & \vdots \\ \pi_{i,N_i,1}\sigma_g^2 & \pi_{i,N_i,2}\sigma_g^2 & \cdots & \sigma_g^2 + \sigma_e^2 \end{pmatrix} \quad (2)$$

where $\pi_{i,j,k}$ is the kinship coefficient between individuals j and k in the i -th family. For example, if individuals j and k are siblings, $\pi_{i,j,k} = 1/4$. It should be noted that if the sample only consists of unrelated individuals, the covariance matrix 2 is degenerated to a scalar. In this case, the variance components g_{ij} and e_{ij} cannot be separately estimated. It is only possible to include the residual variance component $\varepsilon_{ij} = g_{ij} + e_{ij}$ in the model.

The model can be further extended to incorporate additive and dominant effects, shared and independent environments, interaction effects between different components, etc. For simplicity, in this article, we assumed that the environmental components are independent between different family members. This assumption is only valid when shared environment does not contribute to the variability of the QT. However, it is easy to see that the inference for parameter β_1 will not be affected, even when this assumption does not hold. For example, in a sib-pair study, the mixed model is equivalent to $Y_{ij} = \beta_0 + \beta_1 C(\vec{X}_{ij}, Y_{ij}) + \varepsilon_{ij}$. For sib 1 and 2 in sib-pair i , the residual $(\varepsilon_{i,1}, \varepsilon_{i,2})$ follows a bi-variate normal distribution

$$N\left((0,0), \begin{pmatrix} \tau^2 & \rho\tau^2 \\ \rho\tau^2 & \tau^2 \end{pmatrix}\right)$$

with $\tau^2 = \sigma_g^2 + \sigma_e^2$. Fitting the equivalent model does not require explicitly modeling different variance components.

It has been previously demonstrated that it is advantageous to jointly analyze multiple variants within a region in order to improve the power for detecting associations with RV [9, 16–19, 22]. A genotype score is usually assigned for the multi-site genotype $\vec{X}_{ij}$ using the coding function $C(\vec{X}_{ij}, Y_{ij})$. Through defining different coding functions $C(\vec{X}_{ij}, Y_{ij})$, many RV association methods developed for unrelated individuals can be generalized to pedigree samples. Examples of coding functions include

(1) Collapsing Coding: an indicator of whether the individual carries RV is defined, i.e.

$$C^{CMC}(\vec{X}_{ij}, Y_{ij}) = \delta(\sum_{s \in RV} x_{i,j}^s > 0) \quad (3)$$

where δ is an indicator function. RV is the set of RV that are grouped, which can be determined by fixed or variable allele frequency thresholds [9, 17], functional annotations [9, 17], genetic pathways [19], etc. Through the collapsing coding, the CMC and VT tests can be straightforwardly extended in the MEGA framework, where CMC uses fixed and VT uses variable minor allele frequency (MAF) threshold for grouping variants. Statistical significance for CMC can be evaluated analytically, while p values for VT need to be obtained via permutation.

(2) Weighted Sum Coding: variants at each nucleotide site are assigned a weight, and then the weighted genotypes at different nucleotide sites are aggregated [16], i.e.

$$C^{WSC}(\vec{X}_{ij}, Y_{ij}) = \sum_{s \in RV} w^s x_{i,j}^s \quad (4)$$

For detecting QTL associations using related individuals, some possible coding functions include

(2a) ANRV Coding: the coding is determined by the number of RV in the gene locus that are analyzed, i.e.

$$C^{ANRV}(\vec{X}_{ij}, Y_{ij}) = \sum_{s \in RV} x_{i,j}^s \quad (5)$$

(2b) The coding or grouping scheme that uses functional prediction scores or evolutionary information. The weights w^s can be determined by estimated variant frequencies [16], prediction scores from bioinformatics software [17] or based upon estimated fitness values [38]. If the weights are not dependent on the phenotypes, i.e. $C(\vec{X}_{ij}, Y_{ij}) = C(\vec{X}_{ij})$, the statistical significance can be evaluated analytically using asymptotic approximations.

Modeling Selective Sampling Designs

When individuals are ascertained based upon their QTVs, the ascertainment mechanisms need to be adjusted in order to perform valid inferences. Here, a prospective likelihood approach is applied, which models the distribution of QT conditional on the ascertainment mechanism. It uses of all available genotype and phenotype data, and allows efficient inferences for the parameters of interest, such as locus-genetic effects.

The prospective likelihood approach for ascertainment corrections was previously used by Epstein et al. [39] in binary trait linkage analysis. However, the implementation in Epstein et al. [39] is not directly applicable to QT association studies, since the phenotype models, (binary vs. quantitative) as well as the study

designs are different. We extend this method to detect RV QTL associations using unrelated and pedigree samples with extreme QTV. Likelihood models are also derived for more complicated study designs such as ES-ECSP or ES-EP, where only one sib with the most extreme trait value is sequenced from each selected sib-pair.

For studies using pedigree samples, it is assumed that the phenotypes of N_i individuals in pedigree i are involved in sample ascertainment, but only the first N_i^* individuals are sequenced. The joint phenotype distribution of N_i individuals in pedigree i conditional on the sampling status is given by

$$P(Y_{i,1}, \dots, Y_{i,N_i} | \vec{X}_{i,1}, \dots, \vec{X}_{i,N_i}, A_i = 1; \beta_1) \\ = \prod_i \frac{\sum_{\vec{X}_{i,N_i^*+1}, \dots, \vec{X}_{i,N_i}} P(A_i = 1 | Y_{i,1}, \dots, Y_{i,N_i}, \vec{X}_{i,1}, \dots, \vec{X}_{i,N_i}) \times \\ P(Y_{i,1}, \dots, Y_{i,N_i} | \vec{X}_{i,1}, \dots, \vec{X}_{i,N_i}; \beta_1) \times \\ P(\vec{X}_{i,N_i^*+1}, \dots, \vec{X}_{i,N_i} | \vec{X}_{i,1}, \dots, \vec{X}_{i,N_i^*})}{\int P(A_i = 1 | y_{i,1}, \dots, y_{i,N_i}, \vec{X}_{i,1}, \dots, \vec{X}_{i,N_i}) \times \\ dP(y_{i,1}, \dots, y_{i,N_i} | \vec{X}_{i,1}, \dots, \vec{X}_{i,N_i})} \quad (6)$$

where A_i is an indicator of pedigree i being sampled.

Ascertainment mechanisms are modeled through the probability of being sampled, i.e. $P(A_i = 1 | Y_{i,1}, \dots, Y_{i,N_i}, \vec{X}_{i,1}, \dots, \vec{X}_{i,N_i^*})$. Since the sample ascertainment mechanisms discussed in this article only depend on QT, the sampling probability is conditionally independent of genotypes, i.e. $P(A_i = 1 | Y_{i,1}, \dots, Y_{i,N_i}, \vec{X}_{i,1}, \dots, \vec{X}_{i,N_i^*}) = P(A_i = 1 | Y_{i,1}, \dots, Y_{i,N_i})$.

Applications to Study Designs Using Sibs and Unrelated Individuals

Evaluating the likelihood function in equation 6 is generally complicated and computationally intensive. However, for small pedigrees such as sib-pairs, the likelihood configurations can be simplified.

For study designs using EUI, if it is assumed that samples with $QTV > y^{ub}$ or $< y^{lb}$ are collected, the sampling probability satisfies

$$P(A_i = 1 | Y_i) \propto \begin{cases} 1/P(Y_i \in \Omega_{EUI}; \beta_1) & \text{if } Y_i \in \Omega_{EUI} \\ 0 & \text{if } Y_i \notin \Omega_{EUI} \end{cases} \quad (7)$$

if $Y_i \in \Omega_{EUI} = (y^{ub}, \infty) \cup (-\infty, y^{lb})$ if $Y_i \notin \Omega_{EUI}$

The sampling mechanisms for sib-pair study designs are similarly modeled. For instance, in an EP study, each selected sib-pair must contain at least one sib with trait values exceeding the thresholds. The sampling probability thus satisfies

$$P(A_i = 1 | Y_{i,1}, Y_{i,2}) \propto \begin{cases} 1/P((Y_{i,1}, Y_{i,2}) \in \Omega_{EP}; \beta_1) & (Y_{i,1}, Y_{i,2}) \in \Omega_{EP} \\ 0 & (Y_{i,1}, Y_{i,2}) \notin \Omega_{EP} \end{cases} \quad (8)$$

where the constraint Ω_{EP} is given by

$$\Omega_{EP} = \{(Y_1, Y_2) | Y_1 \geq y^{ub}\} \cup \{(Y_1, Y_2) | Y_2 \geq y^{ub}\} \\ \cup \{(Y_1, Y_2) | Y_1 \leq y^{lb}\} \cup \{(Y_1, Y_2) | Y_2 \leq y^{lb}\}$$

Two study designs use phenotype information from both sibs in sample ascertainment, but only sequence one sib from each selected sib-pair, i.e. ES-EP (design 6) and ES-ECSP (design 7). In order to make valid inferences for the genetic parameters of interest, it is necessary to incorporate QT values from the sibs that are not sequenced.

Without loss of generality, it is assumed that sib 1 from each selected sib-pair has the most extreme trait and is sequenced in ES-ECSP and ES-EP. The likelihood for pedigree i can be simplified for study designs using sib-pairs,

$$P(Y_{i,1}, Y_{i,2} | \vec{X}_{i,1}, A_i = 1; \beta_1) \\ = \prod_i \frac{\sum_{\vec{X}_{i,2}} P(A_i = 1 | Y_{i,1}, Y_{i,2}, \vec{X}_{i,1}, \vec{X}_{i,2}) \times \\ P(Y_{i,1}, Y_{i,2} | \vec{X}_{i,1}, \vec{X}_{i,2}; \beta_1, \theta) P(\vec{X}_{i,2} | \vec{X}_{i,1})}{\int P(A_i = 1 | y_{i,1}, y_{i,2}, \vec{X}_{i,1}) \times \\ dP(y_{i,1}, y_{i,2} | \vec{X}_{i,1})} \quad (9)$$

Approximations for $P(Y_{i,1}, Y_{i,2} | \vec{X}_{i,1}, A_i = 1; \beta_1, \theta)$

The exact evaluation for the mixture likelihood equation 9 involves estimating RV frequencies, which is not numerically stable. If the collapsing coding scheme, i.e. $C^{CMC}(\vec{X}_{ij}, Y_{ij}) = \delta(\sum_s x_{ij}^s > 0)$, is used, an approximate likelihood function can be used. Assuming that Hardy-Weinberg equilibrium holds in the parental generations and that parent-offspring allele transmission follows Mendel's law, the probabilities $P(\vec{X}_{i,2} | \vec{X}_{i,1})$ (or $P(C(\vec{X}_{i,2}) | C(\vec{X}_{i,1}))$) can be expressed in terms of the total RV MAFs p in the general population.

According to Bayes' law,

$$P(C(\vec{X}_{i,2}) | C(\vec{X}_{i,1})) = \frac{P(C(\vec{X}_{i,1}), C(\vec{X}_{i,2}))}{P(C(\vec{X}_{i,1}))} \quad (10)$$

An essential part is to compute $P(C(\vec{X}_{i,1}), C(\vec{X}_{i,2}))$. If it is assumed that most locus haplotypes contain no more than 1 RV (compatible with observations from both real and simulated data), the population carrier frequency q can be approximated by $q = 2p(1 - p) + p^2$, where p is the total MAF of RV in the general population. For notational simplicity, a haplotype that contains RV is denoted by M (mutant haplotype), while a haplotype that does not contain rare mutations is denoted by W (wild-type haplotype). Parental mating types and their corresponding probabilities are summarized in table 2. If it is assumed that the de novo mutation rate is low and negligible (arguably true for point mutations within a short gene coding region), the offspring locus genotype probabilities can be calculated according to Mendel's law. The probability $P(C(\vec{X}_{i,1}) = 1, C(\vec{X}_{i,2}) = 1)$ satisfies $P(C(\vec{X}_{i,1}) = 1, C(\vec{X}_{i,2}) = 1) = P(MM \& MM) + P(MW \& MM) + P(MW \& MW)$. $P(MM \& MM)$, $P(MW \& MM)$ and $P(MW \& MW)$ are the probabilities of unordered offspring genotype configurations. The details for computing these probabilities can be calculated from table 2.

When weighted sum coding is used, the genotypes for the sibs that are not sequenced are treated as missing data and replaced by their expected values. Similar strategies have been commonly applied to genetic association studies with imputed data [40]. The approximate likelihood for sib-pair i is thus given by

Table 2. Probabilities of parental mating type and conditional probabilities of sib-pair genotype configurations

	Parental mating type probabilities	Probability of sib-pair genotype configurations conditional on parental genotypes					
		WW&WW	WW&WM	WW&MM	WM&WM	WM&MM	MM&MM
<i>Parental mating type</i>							
WW × WW	$(1 - p)^4$	1	0	0	0	0	0
WW × WM	$4p(1 - p)^3$	1/4	1/2	0	1/4	0	0
WW × MM	$2p^2(1 - p)^2$	0	0	0	1	0	0
WM × WM	$4p^2(1 - p)^2$	1/16	1/4	1/8	1/4	1/4	1/16
WM × MM	$4p^3(1 - p)$	0	0	0	1/4	1/2	1/4
MM × MM	p^4	0	0	0	0	0	1

$$P(Y_{i,1}, Y_{i,2} | \vec{X}_{i,1}, A_i = 1; \beta_1) = \prod_i \frac{P(A_i = 1 | Y_{i,1}, Y_{i,2}, \vec{X}_{i,1}, \hat{\vec{X}}_{i,2}) P(Y_{i,1}, Y_{i,2} | \vec{X}_{i,1}, \hat{\vec{X}}_{i,2}; \beta_1)}{\int P(A_i = 1 | y_{i,1}, y_{i,2}, \vec{X}_{i,1}) dP(y_{i,1}, y_{i,2} | G_{i,1}, \hat{\vec{X}}_{i,2})} \quad (11)$$

where

$$\hat{\vec{X}}_{i,2} = E(\vec{X}_{i,2} | \vec{X}_{i,1}) \text{ and } C^{ANRV}(\hat{\vec{X}}_{i,2}) = \sum_s \hat{X}_{i,2}^s.$$

Association Testing

Hypothesis testing for RV associations can be performed using efficient score statistics. The null hypothesis of no genotype-phenotype association is tested, i.e. $H_0: \beta_1 = 0$. A likelihood-based score test is asymptotically normally distributed and locally most powerful if the model is correctly specified and the genotype coding does not depend on QTs. Other likelihood-based tests such as the likelihood ratio test or Wald test require maximizing the likelihood under the alternative hypothesis. They are not appropriate for detecting associations with RV, since the optimization procedure can be numerically unstable in the presence of sparse data [41].

When data-based weights are applied or when a variable selection procedure is used, normality of the test statistic may not hold. In this case, permutation algorithms are needed to evaluate statistical significance empirically. In addition, p values obtained by asymptotic approximations may not be accurate. It is also necessary in practice to verify significant analytical p values using permutations.

‘Gene dropping’-based permutation was previously developed for a QT transmission disequilibrium test (QTDT) [8, 28]. In a QTDT-based test, the locus genotype coding is decomposed into a within-family and a between-family component, i.e. $C(\vec{X}_i, Y_i) = C_b(\vec{X}_i, Y_i) + C_w(\vec{X}_i, Y_i)$; and the association between the within-family genotype coding $C_w(\vec{X}_i, Y_i)$ and QT is tested [8, 28]. When the founder genotypes are available and it is of interest to detect QT associations using the locus genotype coding, permutations can be carried out by simultaneously shuffling the founder genotypes and performing the gene dropping algorithm.

In this article, permutation algorithms were developed for sib-pair study designs, which sequence both sibs or only one sib with the most extreme trait from each selected sib-pair. The algorithm can be used to empirically obtain statistical significance. Under H_0 of no genotype/phenotype associations, bi-variate genotypes for sib-pair a and sib-pair b are interchangeable, i.e.

$$P(\vec{X}_{a,1}, \vec{X}_{a,2} | Y_{a,1}, Y_{a,2}) = P(\vec{X}_{b,1}, \vec{X}_{b,2} | Y_{b,1}, Y_{b,2}) \\ = P(\vec{X}_{a,1}, \vec{X}_{a,2} | Y_{b,1}, Y_{b,2}) = P(\vec{X}_{b,1}, \vec{X}_{b,2} | Y_{a,1}, Y_{a,2})$$

In addition, within each sib-pair a , the genotypes for the two sibs are exchangeable, i.e.

$$P(\vec{X}_{a,1}, \vec{X}_{a,2} | Y_{a,1}, Y_{a,2}) = P(\vec{X}_{a,2}, \vec{X}_{a,1} | Y_{a,1}, Y_{a,2}) \\ = P(\vec{X}_{a,1}, \vec{X}_{a,1} | Y_{a,1}, Y_{a,2}) = P(\vec{X}_{a,2}, \vec{X}_{a,2} | Y_{a,1}, Y_{a,2})$$

Therefore the following algorithm is motivated, which is used to obtain statistical significance empirically. It is similar to the QFAM test implemented in PLINK [42]. However, it will not induce non-integer genotypes in permuted datasets, which cannot be handled by some RV tests in their implementations.

Algorithm for Study Designs That Sequence Both Sibs from Selected Sib-Pairs

Assume that the dataset consists of N sib-pairs, i.e. sib-pair 1, 2, ..., N , whose genotypes and phenotypes are given by $\{\vec{X}_{1,1}, \vec{X}_{1,2}, Y_{1,1}, Y_{1,2}\}, \dots, \{\vec{X}_{N,1}, \vec{X}_{N,2}, Y_{N,1}, Y_{N,2}\}$.

(1) Calculate the test statistic using the original dataset $\{\vec{X}_{1,1}, \vec{X}_{1,2}, Y_{1,1}, Y_{1,2}\}, \dots, \{\vec{X}_{N,1}, \vec{X}_{N,2}, Y_{N,1}, Y_{N,2}\}$, and denote the statistic as T^{dat} .

Repeat steps 2–5, M times.

For each $m \leq M$,

(2) Shuffle the labels for sib-pairs, i.e. 1, 2, ..., N , and the labels after permutation are denoted by $i^m(1), \dots, i^m(N)$.

(3) Within each sib-pair i , re-sample the labels for sib 1 and 2 with replacement, and the resulting labels are given by $(j^m(1), j^m(2))$, which can be equal to (1,1), (1,2), (2,1), or (2,2).

(4) Put together the permuted dataset m which consists of

$$\left\{ \vec{X}_{i^m(1), j^m(1)}, \vec{X}_{i^m(N), j^m(N)}, Y_{1,1}, Y_{1,2} \right\}, \dots, \\ \left\{ \vec{X}_{i^m(N), j^m(N)}, \vec{X}_{i^m(N), j^m(N)}, Y_{N,1}, Y_{N,2} \right\}$$

(5) Analyze the permuted dataset with MEGA, and the statistic is denoted by T^m .

The empirical p value is obtained by comparing T^{dat} with $\{T^m\}_m$, i.e.

$$\hat{p} = 1/M \sum_{m=1}^M \delta(T^m > T^{dat})$$

The algorithm applies to ECSP (design 2), EDSP (design 3), EDAC (design 4) and EP (design 5).

For the study designs that sequence only one sib from each selected sib-pair, it is assumed that within each sib-pair, sib 1 has the most extreme trait, and is sequenced. The following permutation algorithm can be applied to obtain p values empirically.

Algorithm for Study Designs That Sequence Only One Sib from Each Selected Sib-Pair

Assume that the dataset consists of N sib-pairs, i.e. sib-pair 1, 2, ..., N , whose genotypes and phenotypes are given by $\{\vec{X}_{1,1}, Y_{1,1}, Y_{1,2}\}, \dots, \{\vec{X}_{N,1}, Y_{N,1}, Y_{N,2}\}$.

(1) Calculate the test statistic T^{dat} using the original dataset $\{\vec{X}_{1,1}, Y_{1,1}, Y_{1,2}\}, \dots, \{\vec{X}_{N,1}, Y_{N,1}, Y_{N,2}\}$.

Repeat steps 2–4, M times.

For each $m \leq M$,

(2) Shuffle the labels for sib-pairs, i.e. 1, 2, ..., N , and the labels after permutation are denoted by $i^m(1), \dots, i^m(N)$.

(3) Put together the permuted dataset m which consists of

$$\left\{ \vec{X}_{i^m(1), j_{i^m(1)}^m}, Y_{1,1}, Y_{1,2} \right\}, \dots, \left\{ \vec{X}_{i^m(N), j_{i^m(N)}^m}, Y_{N,1}, Y_{N,2} \right\}$$

(4) Analyze the permuted dataset with MEGA, and the statistic is denoted by T^m .

The empirical p value is obtained by comparing T^{dat} with $\{T^m\}_m$, i.e.

$$\hat{p} = 1/M \sum_{m=1}^M \delta(T^m > T^{dat})$$

The algorithm is applicable to ES-ECSP (design 6) and ES-EP (design 7).

Simulation of Genetic Data

Population-genetic data was generated using forward time simulation [43] incorporating both demographic change and purifying selections. Using parameters estimated by Boyko et al. [34] for African populations, sequence data were simulated. A simple two-epoch model with two degrees of freedom was used to describe the demographic change, where the population was constant with a population effective size of $N_{anc} = 7,778$, followed by a population expansion 6,809 generations ago, before reaching the current effective population size of $N_{curr} = 25,636$. To model purifying selections, the selective disadvantages of new heterozygous and homozygous mutations are assumed to be u and $2u$, respectively. A gamma distribution was chosen to model the scaled selective disadvantage $\gamma = 2N_{curr}u$.

$$\gamma = -r, r \sim \frac{\beta_A^{\alpha_A}}{\Gamma(\alpha_A)} r^{\alpha_A-1} \exp(-\beta_A r)$$

The parameters in the model are given by $\alpha_A = 0.184$, $\beta_A = 8,200$. These demographic and selection models have been shown to be parsimonious and provide good fit to real data. A mutation rate of $\mu_s = 1.8 \times 10^{-8}$ per nucleotide per generation is assumed. On average, the coding region for a human gene is 1,500 base pairs (bp) long [44, 45], therefore 1,500 bp was used in the simulation to specify the locus-scaled mutation rate. One hundred haplotype pools were generated.

To simulate samples for evaluating power and type I errors, a pool was randomly chosen for each replicate. The multi-site genotypes of each unrelated individual and founder ‘individual’ (i.e. parents of sib-pairs) are composed of two randomly chosen haplotypes from the pool. The sib-pair multi-site genotypes are obtained by pairing one paternal and one maternal haplotype. As suggested in Kryukov et al. [46], only non-synonymous (NS) variants are tested in order to reduce the impact of non-causal variants, and increase the signal-to-noise ratio in the study [34].

Simulation of QT

In order to generate QT for individuals, 50% of the nucleotide sites with NS variants are randomly chosen as causal, and influence the QTV. The QTV for individual i is generated according to

$$Y_i \sim N\left(\beta_1 \sum_{s \in S_C} x_i^s, \sigma_r^2\right) \quad (12)$$

where S_C is the set of nucleotide sites containing causal variants.

Similarly, the QTV for sib-pair i are determined by

$$(Y_{i,1}, Y_{i,2}) \sim N(\mu, \Sigma) \quad (13)$$

where

$$\vec{\mu} = \left(\beta_1 \sum_{s \in S_C} x_{i,1}^s, \beta_1 \sum_{s \in S_C} x_{i,2}^s \right)$$

and Σ is given by

$$\Sigma = \begin{pmatrix} \sigma_g^2 + \sigma_e^2 & 1/2 \times \sigma_g^2 \\ 1/2 \times \sigma_g^2 & \sigma_g^2 + \sigma_e^2 \end{pmatrix}$$

According to genetic models 12 and 13, the following genetic parameters can be mathematically defined:

(a) locus-genetic variance:

$$\sigma_G^2 = (\beta_1)^2 \text{var} \left(\delta \left(\sum_{s \in S_C} x_{i,1}^s > 0 \right) \right) \quad (14)$$

(b) total QT variance:

$$\sigma_T^2 = \sigma_G^2 + \sigma_g^2 + \sigma_e^2 \quad (15)$$

(c) residual variance:

$$\sigma_r^2 = \sigma_g^2 + \sigma_e^2 \quad (16)$$

(d) overall heritability:

$$h^2 = \frac{\sigma_G^2 + \sigma_g^2}{\sigma_T^2} \quad (17)$$

(e) locus-specific heritability:

$$h_G^2 = \frac{\sigma_G^2}{\sigma_T^2} \quad (18)$$

(f) sibling-residual correlation:

$$\rho = \frac{2\pi\sigma_g^2}{\sigma_r^2} \quad (19)$$

When evaluating power and type I errors, the residual variance σ_r^2 is fixed to be 1 and for each specified overall heritability value, the corresponding polygenic variance can be calculated according to formulae 15–17.

Type I Error Evaluation

We have evaluated the type I error for the MEGA method using the ANRV coding function, i.e. $C^{ANRV}(\vec{X}_{i,j}, Y_{i,j}) = \sum_s x_{i,j}^s$ [16].

Under H_0 of no association, the distributions of analytical and permutation-based p values of MEGA were examined for different study designs and a variety of overall heritabilities. The empirical p values were estimated based upon 3,000 permutations. For each scenario considered, the empirical distribution of p values was obtained by generating 10,000 replicates under H_0 (i.e. $\beta_1 = 0$). The empirical distributions are plotted against their theoretical expectations.

Power Comparison

Power comparisons are shown when the ANRV coding is used in MEGA. Seven study designs using sib-pairs or unrelated individuals were evaluated using extensive simulations. For cohort selection, in order to compare the power for 6 different sib-pair designs and 1 unrelated individual study design, 2 different causal variant effects and 2 different sample sizes were used, i.e. when $\beta_1 = 1 \sigma_r$, the power of the 6 sib-pair study designs are calculated when 700 selected sib-pairs (or 1 sib per sib-pair for 1,400 selected sib-pairs) are sequenced from a cohort of 7,000 sib-pairs. The power for EUI is calculated when 1,400 samples with extreme traits are sequenced from a cohort of 14,000 unrelated individuals. For a smaller causative variant genetic effect, i.e. $\beta_1 = 0.5 \sigma_r$, the power for sib-pair study designs is examined when 1,600 selected sib-pairs (or 1 sib per sib-pair from 3,200 selected sib-pairs) are sequenced from a cohort of 16,000 sib-pairs. The power for EUI is calculated using 3,200 samples selected from a cohort of 32,000 unrelated individuals.

Second, the power of the unrelated individual study and 6 sib-pair studies was compared for population ascertainment where the selection of sib-pairs (or unrelated individuals) was implemented using pre-specified phenotypic thresholds. The same phenotypic cutoffs $y^{ub} = F^{-1}(80\%)$, $y^{lb} = F^{-1}(20\%)$ (F is the distribution of QTV in the general population) were used for each study design. When $\beta_1 = 0.5 \sigma_r$, 1,600 sib-pairs (or 3,200 unrelated individuals, or 1 sib per sib-pair from 3,200 sib-pairs) were sequenced, and when $\beta_1 = 1 \sigma_r$, 700 sib-pairs (or 1,400 unrelated individuals, or 1 sib per sib-pair from 1,400 sib-pairs) were sequenced. Statistical significance for the test statistic was evaluated analytically. An exome-wide significance level $\alpha = 2.5 \times 10^{-6}$ was used for all power comparisons.

The Analysis of the Ottawa Obesity Study Dataset

Extreme sequencing was carried out for the Ottawa Obesity Study. After controlling for age and sex, 56 genes were sequenced using DNA samples from 378 extremely obese individuals with a BMI above the 95th percentile and 379 lean individuals with a BMI below the 10th percentile. Nineteen of these genes were previously implicated in the etiology of monogenic obesity, i.e. *BRS3*, *CART*, *FABP4*, *HTR2C*, *IL6*, *LEPTIN*, *MC3R*, *MC4R*, *NHLH2*, *NMU*, *NPB*, *GPR7*, *NPY1R*, *NPY2R*, *NPY5R*, *ATGL*, *POMC*, *PYY*, and *UCP3*. The remaining 37 genes are candidate genes for complex polygenic obesity phenotypes, i.e. *ADIPOQ*, *AGRP*, *APOA*, *ARNT2*, *ASIP*, *CIQTNF2*, *C3AR1*, *CCK*, *CPT1B*, *CSF2*, *DGAT1*, *DGAT2*, *GHRL*, *GHSR*, *HSD11B1*, *HTR7*, *INSIG1*, *INSIG2*, *LIPC*, *NMUR1*, *NMUR2*, *NPBWR2*, *NPY*, *NTS*, *PPARGC1A*, *PPY*, *PRKAA1*, *PRKAA2*, *PRKAB1*, *PRKAB2*, *PRKAG1*, *PRKAG2*, *PRKAG3*, *RETN*, *SIRT1*, *TGFBR2*, and *WDTCL1*. Among the 378 extremely obese individuals that were sequenced, 18 were sampled from pedigrees. Phenotype information was available on their related family members and was incorporated in the analysis.

In order to aggregate RV from multiple genes, 19 monogenic obesity-related genes and the 37 candidate genes for complex obesity were respectively grouped and jointly tested. Only NS variants with a frequency $<1\%$ were analyzed. Two analyses were carried out using the extended CMC [9], ANRV [15] and VT [17] methods. In the first analysis, only phenotypes and genotypes from unrelated samples (i.e. 360 obese and 379 lean individuals) were included. In the second analysis, all the sequenced samples were analyzed, which includes both the unrelated samples and the 18 obese individuals selected from pedigrees. In order to model the ascertainment mechanism, quantitative phenotypes from related pedigree members of the 18 obese individuals were also incorporated.

Results

Enrichment of RV

Sample RV frequencies for each study design are shown in table 3. The statistical power of rejecting H_0 of no gene/QT associations is influenced by the enrichment of RV. Based upon the simulated African SFS, RV frequencies in selected samples were examined analytically for the scenario where each causative variant shifts the mean QT value by 1 SD.

Using unrelated individuals or pedigree samples with extreme QTV offers a dramatic increase in aggregated RV frequencies, compared to using randomly ascertained individuals from the general population. RV frequencies can be further increased if pedigrees with multiple members having extreme traits are sequenced. For example, for population ascertainment, when the overall heritability is 10%, the RV carrier frequencies in the samples from EUI and ECSP are 3.5 and 5.2%, respectively, which displays a great enrichment of RV compared to the population carrier frequency of 1.8%.

For study designs using equal-sized cohorts of unrelated individuals and sib-pairs, RV are most enriched in the samples from EP (design 5) and EUI (design 1). For example, when the overall heritability is 10%, the RV frequencies in EUI and EP samples are 5.3 and 5.4%, respectively. When the overall heritability is 90%, the RV carrier frequency in EP samples is slightly decreased (5.2%).

Type I Error Evaluation

The type I errors of MEGA for 7 study designs were evaluated for scenarios with different overall heritabilities. The empirical distributions of analytical p values (fig. 2) and p values obtained via permutations (fig. 3) are plotted for the scenarios where population ascertainment was implemented and the overall heritability is 10 and

Table 3. RV frequencies in selected samples

Study design	Mathematical description	Carrier frequencies Overall heritability				
		10%	30%	50%	70%	90%
Cohort selection						
1. EUI	$P\left(C\left(\vec{X}\right)=1\middle Y\in\Omega_{EUI}\right)$	0.053	0.053	0.053	0.053	0.053
2. ECSP	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{ECSP}\right)$	0.050	0.049	0.047	0.046	0.045
3. EDSP	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{EDSP}\right)$	0.036	0.038	0.039	0.041	0.043
4. EDAC	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{EDAC}\right)$	0.052	0.052	0.051	0.050	0.048
5. EP	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{EP}\right)$	0.054	0.053	0.053	0.053	0.052
6. ES-ECSP	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{ES-ECSP}\right)$	0.043	0.042	0.042	0.041	0.041
7. ES-EP	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{ES-EP}\right)$	0.051	0.050	0.050	0.050	0.050
Population ascertainment						
General population	$P\left(C\left(\vec{X}\right)=1\right)$	0.018	0.018	0.018	0.018	0.018
1. EUI	$P\left(C\left(\vec{X}\right)=1\middle Y\in\Omega_{EUI}\right)$	0.035	0.035	0.035	0.035	0.035
2. ECSP	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{ECSP}\right)$	0.052	0.048	0.045	0.042	0.039
3. EDSP	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{EDSP}\right)$	0.038	0.041	0.045	0.051	0.060
4. EDAC	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{EDAC}\right)$	0.046	0.046	0.045	0.043	0.042
5. EP	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{EP}\right)$	0.031	0.031	0.031	0.031	0.031
6. ES-ECSP	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{ES-ECSP}\right)$	0.056	0.051	0.048	0.045	0.043
7. ES-EP	$P\left(C\left(\vec{X}_1\right)=1\middle (Y_1,Y_2)\in\Omega_{ES-EP}\right)$	0.033	0.033	0.033	0.034	0.034
For cohort selections, 10% of the samples selected from the extremes are sequenced for each design. For population ascertainment, samples are ascertained using thresholds $y^{\mu b}=F^{-1}$ (80%) and $y^{lb}=F^{-1}$ (20%). The causative variant genetic effect of $\beta_1=1$ σ_r is assumed.

Fig. 2. Quantile-quantile plots of analytic p values obtained using MEGA under H_0 against their theoretical expectations. Results for the analysis of the 7 study designs are shown for 2 different overall heritabilities: $h^2 = 10\%$ and $h^2 = 90\%$. A total of 1,400 se-

lected individuals (1,400 unrelated individuals or 700 sib-pairs, or 1 sib per sib-pair for 1,400 sib-pairs) were used for each study design. 10,000 replicates were generated for each study design/overall heritability combination.

90%. They match well with their theoretical expectations under H_0 . The null distributions of p values in other scenarios with different overall heritabilities (i.e. h^2 is 30, 50 and 70%) were also investigated and the type I error was well controlled (data not shown).

Power Comparisons of Sib-Pair and Unrelated Individual Study Designs under Cohort Selections

One unrelated individual study design and 6 sib-pair study designs are compared when selective sampling is carried out within an existing cohort of sib-pairs or an equal-sized cohort of unrelated individuals (fig. 4). The power for the 7 study designs largely follows in the order of EUI (design 1) > ES-EP (design 7) $\geq$ EP (design 5) > EDAC (design 4) > ECSP (design 2) $\approx$ ES-ECSP (design 6) > EDSP (design 3) (see fig. 1 and table 1 for detailed definitions for each study design).

The power for EUI is consistently the highest. The advantage over the sib-pair study designs increases with trait residual correlations. For example, when the effect of causal variant is to shift the mean trait value by 0.5 SD, i.e. $\beta_1 = 0.5 \sigma_r$ and the overall heritability is 10%, the power for EUI is 68.5%. It is slightly higher than that of ES-EP (64.7%), where only the sib with the most extreme trait from each sib-pair selected by proband is sequenced.

However, when the overall heritability is 90%, the power for ES-EP is only 58.2%.

It is very interesting to note that for the studies that use cohorts of sib-pairs, sequencing unrelated samples is still the most powerful strategy. Specifically, ES-EP that sequences the sib with the most extreme trait from each sib-pair selected by proband has the highest power. They are particularly advantageous to the other sib-pair designs when the trait residual correlations are low. For example, when the overall heritability is 10% and the causal variant effect is to shift the mean QTV by 0.5 SD, the power for ES-EP (design 7) is 64.1%. It is much more powerful than the second ranked EP, where both sibs in a selected sib-pair are sequenced (53.2%).

When samples are selected within a finite-sized cohort, the sib-pair study designs can be underpowered if both sibs are required to have extreme trait values. This is because QT varies in a continuous scale, and a more stringent selection criterion will result in less extreme phenotypic thresholds (table 4). For example, for ES-ECSP (design 6) that sequences 1 sib per extreme concordant sib-pair, when the overall heritability is 50%, the upper selection threshold $y_{ES-ECSP}^{ub}$ is 0.620 SD. This cutoff is much less extreme than that for ES-EP (design 7; $y_{ES-EP}^{ub} = 1.612$ SD), where only the sib with the most ex-

Fig. 3. Quantile-quantile plots of empirical distributions of p values obtained via permutations using MEGA under H_0 against their theoretical expectations. Results for the analysis of the 7 study designs are shown for 2 different overall heritabilities: $h^2 = 10\%$ and $h^2 = 90\%$. A total of 1,400 selected individuals

(1,400 unrelated individuals or 700 sib-pairs, or 1 sib per sib-pair for 1,400 sib-pairs) were used for each study design. The p values were obtained using 3,000 permutations. 10,000 replicates were generated for each study design/overall heritability combination.

Fig. 4. Power comparisons for MEGA when selection is implemented from an existing cohort. Power is shown for the 6 sib-pair study designs and 1 unrelated individual design under a variety of causative variant genetic effect/overall heritability combinations. Two causative variant genetic effects ($\beta_1 = 0.5 \sigma_r$ and $\beta_1 = 1 \sigma_r$) together with 5 overall heritabilities ranging from 10 to 90% were examined. For sib-pair study designs, when $\beta_1 = 0.5 \sigma_r$ (a), 1,600 selected sib-pairs (or 1 sib per sib-pair for 3,200 selected sib-pairs) were sequenced from a cohort of 16,000 sib-pairs and, when

$\beta_1 = 1 \sigma_r$ (b), 700 selected sib-pairs (or 1 sib per sib-pair for 1,400 selected sib-pairs) were sequenced from a cohort of 7,000 sib-pairs. For the study design using unrelated individuals, when $\beta_1 = 0.5 \sigma_r$, 3,200 samples with extreme traits were sequenced from a cohort of 32,000 unrelated individuals and, when $\beta_1 = 1 \sigma_r$, 1,400 samples with extreme traits were sequenced from a cohort of 14,000 unrelated individuals. The p values were obtained analytically. For each scenario, the power was evaluated for an exome-wide significance level $\alpha = 2.5 \times 10^{-6}$ using 2,000 replicates.

treme QT is sequenced from each sib-pair selected by proband.

The power of study designs using pedigree samples varies with trait-residual correlations (fig. 4). When the locus-genetic effect is fixed, sib-pair residual correlation increases with overall heritability. The most dramatic trend is observed for EDSP (design 3): with increasing levels of sib-pair residual correlations, the power is increasing. The power for study designs using extreme concordant sib-pairs (i.e. design 2: ECSP and design 6: ES-ECSP) is comparable, and decreases with increasing residual correlations.

Power Comparisons of Sib-Pair and Unrelated

Individual Designs under Population Ascertainment

In practice, it may be necessary to collect and phenotype new samples for studying a novel trait or performing genetic research in a new population. The power for the 6 sib-pair study designs and 1 unrelated individual design was compared when samples are selected from a population of unlimited size using pre-specified phenotypic thresholds (fig. 5). The results are very different from cohort selections. The efficiencies for sib-pair study designs are impacted by the phenotypic residual correlation ρ . When residual correlation is low, the power from greatest to lowest for the 7 study designs are ES-ECSP (design 6) > ECSP (design 2) > EDAC (design 4) $\approx$ ES-EP (design 7) > EUI (design 1) > EP (design 5) $\approx$ EDSP (design 3) (see fig. 1 and table 1 for detailed descriptions of each study design). On the other hand, for QT with high sib-pair residual correlations, the order follows ES-ECSP (design 6) $\geq$ EDSP (design 3) > EDAC (design 4) > ECSP (design 2) $\approx$ ES-EP (design 7) > EUI (design 1) > EP (design 5). The results are largely unaffected by different choices of causal variant genetic effects. Unlike cohort selections, when population ascertainment is performed, proband selection strategies (i.e. design 7: ES-EP and design 5: EP) are no longer among the most powerful designs.

The ES-ECSP (design 6), where only the sib with the most extreme trait value is sequenced from each extreme concordant sib-pair, consistently outperforms other study designs. Its power advantage is particularly pronounced for traits with low residual correlations. However, for a given QT threshold, ES-ECSP requires phenotyping and screening a huge number of sib-pairs. For example, when the trait overall heritability is 10%, approximately 15,913 sib-pairs from the general population will need to be screened in order to obtain 1,400 ECSPs, where both sibs are required to have concordant QTV from the upper or lower 20% extremes.

Table 4. The upper phenotypic cutoff (y^{ub}) for sib-pair study designs using cohort selection (in standard deviations)

Overall heritability	EDSP ^a	ECSP ^a	EDAC ^a	EP ^a	ES-ECSP ^b	ES-EP ^b
10%	0.726	0.793	1.003	1.952	0.506	1.629
30%	0.659	0.860	1.011	1.948	0.563	1.622
50%	0.590	0.927	1.027	1.942	0.620	1.612
70%	0.520	0.995	1.054	1.934	0.679	1.600
90%	0.448	1.064	1.092	1.923	0.740	1.585

^a 1,600 sib-pairs are selected from a cohort of 16,000 sib-pairs.

^b 3,200 sib-pairs are selected from a cohort of 16,000 sib-pairs.

EDAC (design 4) and ECSP (design 2) are among the second most powerful designs, and they are both more powerful than the EUI study design that uses unrelated individuals. EDAC is slightly less powerful for traits with low residual correlations, but has better power than ECSP for traits with high residual correlations. EDAC and ECSP are both consistently more powerful than the EUI. It should be noted that EDAC only screens half as many sib-pairs compared to ECSP, when the overall heritability is 10% (table 5).

The power of EDSP is highly dependent on sib-pair residual correlations. For QT with high residual correlations, the extreme discordant sib-pair design is among the most powerful study designs. However, similar to ES-ECSP, EDSP also requires screening a large number of sib-pairs (table 5). For a trait with an overall heritability of 90%, approximately 32,458 sib-pairs are needed in order to obtain 700 selected sib-pairs. The number of individuals (or sib-pairs) that have to be screened for other study designs can also be found in table 5.

The Analysis of the Ottawa Obesity Study Dataset

A total of 272 variant nucleotide sites were uncovered in the sequence sample from the Ottawa Obesity Study. The dataset is enriched with RV, where >80% of the variants are observed <10 times in the sample.

In the first analysis that uses only unrelated samples, the p values are $p^{CMC} = 0.045$, $p^{ANRV} = 0.037$, and $p^{VT} = 0.035$, which are nominally significant. In the second analysis that uses all available samples (both unrelated individuals and samples selected from pedigrees), smaller p values are observed for all tests ($p^{CMC} = 0.035$, $p^{ANRV} = 0.026$, $p^{VT} = 0.030$). The observation of smaller p values could have occurred due to incorporating data from additional samples, which can improve the power

Fig. 5. Power comparisons for MEGA when samples are ascertained from the general population using fixed QT cutoffs. Power is shown for the 6 sib-pair study designs and 1 unrelated individual design when selection was carried out. Phenotypic thresholds equal to the 20th and 80th percentiles of the trait distribution were used. Power was compared under a variety of causative variant genetic effect/overall heritability combinations. Two causative variant genetic effects ($\beta_1 = 0.5 \sigma_r$ and $\beta_1 = 1 \sigma_r$) together with 5

overall heritabilities ranging from 10 to 90% were examined. For $\beta_1 = 0.5 \sigma_r$ (a), 1,600 sib-pairs (or 1 sib per sib-pair for 3,200 selected sib-pairs, or 3,200 selected unrelated individuals) were analyzed and, for $\beta_1 = 1 \sigma_r$ (b), 700 sib-pairs (or 1 sib per sib-pair for 1,400 selected sib-pairs, or 1,400 selected unrelated individuals) were sequenced. The p values were obtained analytically. For each scenario, the power was evaluated for an exome-wide significance level $\alpha = 2.5 \times 10^{-6}$ using 2,000 replicates.

Table 5. Number of individuals screened for population ascertainment

Study design	Overall heritability				
	10%	30%	50%	70%	90%
EUI ^a	3,500.0	3,500.0	3,500.0	3,500.0	3,500.0
ECSP ^b	7,956.6	6,680.5	5,700.7	4,924.3	4,291.6
EDSP ^b	9,681.8	12,123.2	15,786.2	21,725.5	32,458.4
EDAC ^b	4,367.4	4,307.1	4,188.2	4,014.4	3,790.4
EP ^b	1,094.2	1,098.1	1,106.1	1,118.9	1,137.6
ES-ECSP ^b	15,913.2	13,361.0	11,401.3	9,848.5	8,583.2
ES-EP ^b	1,966.2	2,013.8	2,067.3	2,128.2	2,198.2

The average number of unrelated individuals or sib-pairs screened in order to ascertain 700 sib-pairs, or 1,400 unrelated individuals, or 1 sib per sib-pair for 1,400 selected sib-pairs. Threshold levels $y^{ub} = F^{-1}(80\%)$ and $y^{lb} = F^{-1}(20\%)$ are used for selecting sib-pairs or unrelated individuals under each study design. The causative variant genetic effect of $\beta_1 = 1 \sigma_r$ is assumed.

^a Average number of unrelated individuals screened.

^b Average number of sib-pairs screened.

for detecting QT associations (table 6). The associations with 37 complex obesity candidate genes were also analyzed. However, no significance was identified, which is in concordance with the original study [2].

Discussion

In this article the mixed effects likelihood framework MEGA is introduced for direct association mapping of rare QTL variants. The mixed likelihood framework analyzes QT information, and also allows for efficient inferences of the genetic parameters of interest. Appropriate permutation algorithms were developed to evaluate statistical significance empirically. As a result, a number of permutation-based RV tests developed for unrelated samples can be extended to sib-pairs.

As an application of the MEGA model, we investigated optimal study designs using sib-pairs and unrelated individuals. The results provide 3 important implications for implementing sequence-based genetic studies:

Table 6. Results (p values) for analyzing the sequence dataset from the Ottawa Obesity Study

	Monogenic obesity genes ^a	Complex obesity genes ^b
QT using sequenced unrelated individuals only ^c		
CMC ^d	0.045*	0.635
ANRV ^e	0.037*	0.661
VT ^f	0.035*	0.676
QT using sequenced unrelated individuals and pedigree samples ^c		
CMC ^d	0.035*	0.578
ANRV ^e	0.026*	0.614
VT ^f	0.030*	0.593

^a Nineteen genes that were implicated in monogenic obesity, i.e. *BRS3*, *CART*, *FABP4*, *HTR2C*, *IL6*, *LEPTIN*, *MC3R*, *MC4R*, *NHLH2*, *NMU*, *NPB*, *GPR7*, *NPY1R*, *NPY2R*, *NPY5R*, *ATGL*, *POMC*, *PYY*, and *UCP3*.

^b Thirty-seven candidate genes for complex obesity phenotypes, i.e. *ADIPOQ*, *AGRP*, *APOA*, *ARNT2*, *ASIP*, *CIQTNF2*, *C3AR1*, *CCK*, *CPT1B*, *CSF2*, *DGAT1*, *DGAT2*, *GHRL*, *GHSR*, *HSD11B1*, *HTR7*, *INSIG1*, *INSIG2*, *LIPC*, *NMUR1*, *NMUR2*, *NPBWR2*, *NPY*, *NTS*, *PPARGC1A*, *PPY*, *PRKAA1*, *PRKAA2*, *PRKAB1*, *PRKAB2*, *PRKAG1*, *PRKAG2*, *PRKAG3*, *RETN*, *SIRT1*, *TGFB2*, and *WDTCL*.

^c p values were obtained empirically through 5,000 permutations.

^d Statistical analysis using the collapsing coding method (CMC).

^e Statistical analysis using the coding of the ANRV.

^f Statistical analysis using the VT method.

- (1) Most current genetic studies use existing cohorts of samples. Available cohorts of unrelated individuals can be much larger in size than those of pedigree samples. To implement cohort selections, sequencing samples with extreme traits from a cohort of unrelated individuals can be more powerful than studies that use equal sized or smaller cohorts of sib-pairs.
- (2) The power for sib-pair study designs was also evaluated. Compared to the analyses by Abecasis et al. [8, 27, 28], we directly analyzed the associations between multiple RV in the gene locus and QT, and also considered studies where only 1 sib per selected sib-pair is analyzed. We showed that ES-EP, which sequences the sib with the most extreme trait from each sib-pair selected by proband, outperforms other study designs in power. In addition, as it sequences only one sib from each sib-pair, for a given sequence sample size, genetic variants from a larger number of independent chromosomes can be uncovered. Therefore, ES-EP is also advantageous for discovering and cataloguing novel variants that are of potential population-genetic or medical importance.
- (3) When new samples are ascertained from the population using a pre-specified cutoff, it is important to collect phenotypic information (e.g. systolic and diastolic blood pressure) from additional family members. These information can be easy to obtain and will be

greatly useful for prioritizing samples and increasing power. When QTVs from related pedigree members are used in sample ascertainment, they need to be incorporated in the analysis for valid inferences. The likelihood-based methods described for ES-ECSP (design 6) and ES-EP (design 7) can be applied.

In addition to the sib-pair study designs that were discussed in this article, Kwan et al. [47] proposed an informativeness index for association mapping, and they showed that sequencing individuals with extreme informativeness indices can be more powerful than the 6 sib-pairs study designs discussed in this article. However, the informativeness measure in Kwan et al. [47] cannot be computed for sequence-based association studies. This is because calculating the informativeness index requires enumerating all possible gene locus genotypes and their frequencies, but in sequence-based studies, potential variant sites are not known in advance, and the estimates of RV frequencies are also not reliable. Therefore, the study design proposed in Kwan et al. [47] is not applicable for detecting associations with RV.

The sequence dataset generated by the Ottawa Obesity Study was analyzed. Nominally significant associations were identified between the BMI and the set of 19 monogenic obesity candidate genes. The results confirmed previous analyses [2]. Applying MEGA, pedigree samples and unrelated individuals with extreme traits

can be jointly analyzed. Compared to only analyzing unrelated samples with extreme traits, the power for the joint analysis can be increased and more significant results can be obtained. No significant association was found between the BMI and the set of 37 complex obesity candidate genes. This could have occurred due to (1) the small sample sizes that were analyzed, (2) the moderate effect sizes of RV involved in CT etiologies, (3) some of the 37 genes not being associated with the BMI or (4) variants in the same gene having effects in opposite directions, which negatively impacts power.

Correcting for ascertainment mechanisms has been well researched [48]. In addition to prospective likelihood-based approaches, retrospective [47] and conditional likelihoods [49] were also applied to adjust for sampling ascertainment mechanisms in linkage and CV association studies. However, they are not applicable for detecting RV QT associations. Specifically, conditional likelihood considers the QT distribution of individuals who are not probands conditional on that of the probands. It cannot be applied to studies where QTV of both sibs in a sib-pair are used in sample ascertainment. In retrospective likelihoods, genotypes are treated as random variables and the probability of genotype configurations conditional on QTV is modeled. Computing retrospective likelihoods requires enumerating all possible multi-site genotype configurations and their frequencies, which is not feasible since the frequencies of rare multi-site genotypes cannot be reliably estimated.

Joint modeling of linkage and association using a variance components model has been widely applied to QTL mapping. The model requires calculating identity-by-descent coefficients, which can be computationally intensive for large pedigrees (or inbred populations) or dense sets of markers (especially for sequence data). On the other hand, direct association mapping only requires kinship coefficients, which are easy to compute.

For evaluating different study designs, a set dominance model was used where all causal variants are assumed to have equal effect. A handful of CT studies support our choice of causative variant genetic effect values. For example, it has been observed that a NS variant A390P in *CETP* is associated with a 0.4 SD decrease in plasma high-density lipoprotein cholesterol levels [50].

The impact of different population-genetic and phenotypic models on power comparisons was evaluated. Specifically, we examined the genetic models for which variants of lower frequencies have larger phenotypic effects. The model is motivated by the observation that an inverse relationship may exist between variants' MAFs

and their genetic effects [1]. We also examined the phenotypic model described by Kryukov et al. [46], where the QT is assumed to be visible to purifying selections, and the causality of a variant is determined by its fitness. In this model, variants with selection coefficients greater than a threshold are deemed causal and affect the QT of interest. Recently, it was also suggested that (rare) variants in the same genetic locus may have effects in opposite directions [21, 51, 52]. The power for different RV tests varies under different alternative models. However, the validity of these tests is not affected and the relative efficiency for different study designs remains unchanged under these alternative models (data not shown).

The power comparisons in this article were carried out using an exome-wide significance level $\alpha = 2.5 \times 10^{-6}$, which is based upon a Bonferroni correction for testing 20,000 genes. For exome sequencing, since the analysis is carried out on a gene level and variants in different genes are only very weakly correlated, it is not overly conservative to use Bonferroni corrections [9, 16]. It should be noted that correctly controlling for the family-wise error rate is not sufficient for eliminating spurious associations. It is also important to replicate the identified associations using an independent dataset [53].

For the power comparisons of study designs using pre-specified cutoffs, the impact of different cutoff values y^{ub} , y^{lb} was examined. For the comparisons of selection strategies implemented from within an existing cohort, multiple cohort sizes and selection proportions were also studied. The relative performance of different sampling strategies remains similar. In practice, it is of great interest to know what is the optimal phenotypic threshold for population ascertainment. A universally valid answer to this question is not available. When quantities of clinical interest are available, they can naturally be used as cutoffs. The cutoffs can also be affected by the costs of screening and sequencing individuals. For sequencing a fixed number of individuals, using a more stringent threshold will increase power, but this will also come at the cost of having to screen many more individuals. The use of too stringent thresholds for some phenotypes may not be advisable [54], since extreme outliers can be due to measurement errors or other artifacts [54, 55]. When selection is implemented within an existing cohort, the cutoffs are jointly determined by the QT distribution, the size of the cohort, and the number of individuals to be sequenced. The cohort sizes used in the simulations are representative of existing cohorts, such as the Mid-Atlantic Twin Registry [10] or the UK Adult Twin Registry [56].

With the rapid development of next-generation sequencing technologies, direct exome-wide (or whole genome-wide) association studies will be applied on a much larger scale for mapping CT. The methods and results presented in this article will play an important role in dissecting RVs' CT etiologies.

Acknowledgements

This research is supported by the National Institutes of Health grants HL102926-0110 and MD005964 (to S.M.L.). D.J.L. was partially supported by a training fellowship from the Keck Center Pharmacoinformatics Training Program of the Gulf Coast Consortia (NIH grant No. 5 R90 DK071505-04). We would like to

thank Drs. Bingshan Li and Gao Wang for helpful discussions. We would also like to thank Drs. Ruth McPherson and Robert Dent for providing us with sequence data on 56 obesity genes from the Ottawa Obesity Study, which was supported by a grant from the Canadian Institutes of Health Research MOP-111107 (to Ruth McPherson and Robert Dent). Computation for this research was supported in part by the Shared University Grid at Rice funded by NSF under grant EIA-0216467, and a partnership between Rice University, Sun Microsystems and Sigma Solutions, Inc.

Financial Interest

The authors declare no competing interests.

References

- Bodmer W, Bonilla C: Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* 2008;40:695–701.
- Ahituv N, Kavaslar N, Schackwitz W, Ustaszewska A, Martin J, Hebert S, Doelle H, Ersoy B, Kryukov G, Schmidt S, Yosef N, Ruppin E, Sharan R, Vaisse C, Sunyaev S, Dent R, Cohen J, McPherson R, Pennacchio LA: Medical sequencing at the extremes of human body mass. *Am J Hum Genet* 2007;80:779–791.
- Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH: Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* 2004;305:869–872.
- Cohen JC, Pertsemlidis A, Fahmi S, Esmail S, Vega GL, Grundy SM, Hobbs HH: Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels. *Proc Natl Acad Sci USA* 2006;103:1810–1815.
- Ji W, Foo JN, O'Roak BJ, Zhao H, Larson MG, Simon DB, Newton-Cheh C, State MW, Levy D, Lifton RP: Rare independent mutations in renal salt handling genes contribute to blood pressure variation. *Nat Genet* 2008;40:592–599.
- Romeo S, Pennacchio LA, Fu Y, Boerwinkle E, Tybjaerg-Hansen A, Hobbs HH, Cohen JC: Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. *Nat Genet* 2007;39:513–516.
- Romeo S, Yin W, Kozlitina J, Pennacchio LA, Boerwinkle E, Hobbs HH, Cohen JC: Rare loss-of-function mutations in ANGPTL family members contribute to plasma triglyceride levels in humans. *J Clin Invest* 2009;119:70–79.
- Abecasis GR, Cardon LR, Cookson WO, Sham PC, Cherny SS: Association analysis in a variance components framework. *Genet Epidemiol* 2001;21(suppl 1):S341–S346.
- Li B, Leal SM: Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet* 2008;83:311–321.
- Anderson LS, Beverly WT, Corey LA, Murrelle L: The Mid-Atlantic Twin Registry. *Twin Res* 2002;5:449–455.
- Victor RG, Haley RW, Willett DL, Peshock RM, Vaeth PC, Leonard D, Basit M, Cooper RS, Iannacchione VG, Visscher WA, Staab JM, Hobbs HH: The Dallas Heart Study: a population-based probability sample for the multidisciplinary study of ethnic differences in cardiovascular health. *Am J Cardiol* 2004;93:1473–1480.
- Kagan A, Dawber TR, Kannel WB, Revotskie N: The Framingham study: a prospective study of coronary heart disease. *Fed Proc* 1962;21(Pt 2):52–57.
- Peng B, Li B, Han Y, Amos CI: Power analysis for case-control association studies of samples with known family histories. *Hum Genet* 2010;127:699–704.
- Li M, Boehnke M, Abecasis GR: Efficient study designs for test of genetic association using sibship data and unrelated cases and controls. *Am J Hum Genet* 2006;78:778–792.
- Morris AP, Zeggini E: An evaluation of statistical approaches to rare variant analysis in genetic association studies. *Genet Epidemiol* 2010;34:188–193.
- Madsen BE, Browning SR: A groupwise association test for rare mutations using a weighted sum statistic. *PLoS Genet* 2009;5:e1000384.
- Price AL, Kryukov GV, de Bakker PI, Purcell SM, Staples J, Wei LJ, Sunyaev SR: Pooled association tests for rare variants in exon-resequencing studies. *Am J Hum Genet* 2010;86:832–838.
- Bhatia G, Bansal V, Harismendy O, Schork NJ, Topol EJ, Frazer K, Bafna V: A covering method for detecting genetic associations between rare variants and common phenotypes. *PLoS Comput Biol* 2010;6:e1000954.
- Liu DJ, Leal SM: A novel adaptive method for the analysis of next-generation sequencing data to detect complex trait associations with rare variants due to gene main effects and interactions. *PLoS Genet* 2010;6:e1001156.
- Ionita-Laza I, Buxbaum JD, Laird NM, Lange C: A new testing strategy to identify rare variants with either risk or protective effect on disease. *PLoS Genet* 2010;7:e1001289.
- Neale BM, Rivas MA, Voight BF, Altshuler D, Devlin B, Orho-Melander M, Kathiresan S, Purcell SM, Roeder K, Daly MJ: Testing for an unusual distribution of rare variants. *PLoS Genet* 2011;7:e1001322.
- Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X: Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet* 2011;89:82–93.
- Darvasi A: Closing in on complex traits. *Nat Genet* 2006;38:861–862.
- Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E: Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 2010;42:348–354.
- Zhu X, Feng T, Li Y, Lu Q, Elston RC: Detecting rare variants for complex traits using family and unrelated data. *Genet Epidemiol* 2010;34:171–187.

- 26 Feng T, Elston RC, Zhu X: Detecting rare and common variants for complex traits: sibpair and odds ratio weighted sum statistics (SPWSS, ORWSS). *Genet Epidemiol* 2011;35:398–409.
- 27 Abecasis GR, Cookson WO, Cardon LR: The power to detect linkage disequilibrium with quantitative traits in selected samples. *Am J Hum Genet* 2001;68:1463–1474.
- 28 Abecasis GR, Cardon LR, Cookson WO: A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 2000;66:279–292.
- 29 Risch NJ, Zhang H: Mapping quantitative trait loci with extreme discordant sib pairs: sampling considerations. *Am J Hum Genet* 1996;58:836–843.
- 30 Risch N, Zhang H: Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science* 1995;268:1584–1589.
- 31 Carey G, Williamson J: Linkage analysis of quantitative traits: increased power by using selected samples. *Am J Hum Genet* 1991;49:786–796.
- 32 Eaves L, Meyer J: Locating human quantitative trait loci: guidelines for the selection of sibling pairs for genotyping. *Behav Genet* 1994;24:443–455.
- 33 Gu C, Rao DC: A linkage strategy for detection of human quantitative-trait loci. I. Generalized relative risk ratios and power of sib pairs with extreme trait values. *Am J Hum Genet* 1997;61:200–210.
- 34 Boyko AR, Williamson SH, Indap AR, Degehardt JD, Hernandez RD, Lohmueller KE, Adams MD, Schmidt S, Sninsky JJ, Sunyaev SR, White TJ, Nielsen R, Clark AG, Bustamante CD: Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet* 2008;4:e1000083.
- 35 Falconer DS, Mackay TFC: *Introduction to Quantitative Genetics*, ed 4. Essex, Longman, 1996.
- 36 Fisher RA: The correlation between relatives on the supposition of mendelian inheritance. *Philos Trans R Soc Edinburgh* 1918;52:399–433.
- 37 Chen WM, Abecasis GR: Family-based association tests for genomewide association scans. *Am J Hum Genet* 2007;81:913–926.
- 38 King CR, Rathouz PJ, Nicolae DL: An evolutionary framework for association testing in resequencing studies. *PLoS Genet* 2010;6:e1001202.
- 39 Epstein MP, Lin X, Boehnke M: Ascertainment-adjusted parameter estimates revisited. *Am J Hum Genet* 2002;70:886–895.
- 40 Li Y, Willer C, Sanna S, Abecasis G: Genotype imputation. *Annu Rev Genomics Hum Genet* 2009;10:387–406.
- 41 Mehta CR, Patel NR: Exact logistic regression: theory and examples. *Stat Med* 1995;14:2143–2160.
- 42 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: Plink: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–575.
- 43 Hernandez RD: A flexible forward simulator for populations subject to selection and demography. *Bioinformatics* 2008;24:2786–2787.
- 44 Pritchard JK: Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet* 2001;69:124–137.
- 45 Eyre-Walker A, Keightley PD: High genomic deleterious mutation rates in hominids. *Nature* 1999;397:344–347.
- 46 Kryukov GV, Shpunt A, Stamatoyannopoulos JA, Sunyaev SR: Power of deep, all-exon resequencing for discovery of human trait genes. *Proc Natl Acad Sci USA* 2009;106:3871–3876.
- 47 Kwan JS, Cherny SS, Kung AW, Sham PC: Novel sib pair selection strategy increases power in quantitative association analysis. *Behav Genet* 2009;39:571–579.
- 48 Kraft P, Thomas DC: Bias and efficiency in family-based gene-characterization studies: conditional, prospective, retrospective, and joint likelihoods. *Am J Hum Genet* 2000;66:1119–1131.
- 49 Sung YJ, Dawson G, Munson J, Estes A, Schellenberg GD, Wijsman EM: Genetic investigation of quantitative traits related to autism: use of multivariate polygenic models with ascertainment adjustment. *Am J Hum Genet* 2005;76:68–81.
- 50 Spirin V, Schmidt S, Pertsemlidis A, Cooper RS, Cohen JC, Sunyaev SR: Common single-nucleotide polymorphisms act in concert to affect plasma levels of high-density lipoprotein cholesterol. *Am J Hum Genet* 2007;81:1298–1303.
- 51 Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH: Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet* 2005;37:161–165.
- 52 Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH: Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med* 2006;354:1264–1272.
- 53 Liu DJ, Leal SM: Replication strategies for rare variant complex trait association studies via next-generation sequencing. *Am J Hum Genet* 2010;87:790–801.
- 54 Lander ES, Botstein D: Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 1989;121:185–199.
- 55 Slatkin M: Disequilibrium mapping of a quantitative-trait locus in an expanding population. *Am J Hum Genet* 1999;64:1764–1772.
- 56 Spector TD, Williams FM: The UK Adult Twin Registry (TwinsUK). *Twin Res Hum Genet* 2006;9:899–906.