



Review

# Statistical methods for SNP heritability estimation and partition: A review



Huanhuan Zhu <sup>a</sup>, Xiang Zhou <sup>a,b,\*</sup>

<sup>a</sup> Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA  
<sup>b</sup> Center for Statistical Genetics, University of Michigan, Ann Arbor, MI 48109, USA

ARTICLE INFO

**Article history:**  
 Received 28 February 2020  
 Received in revised form 3 June 2020  
 Accepted 7 June 2020  
 Available online 18 June 2020

**Keywords:**  
 SNP heritability  
 Linear mixed model  
 REML  
 Method of moments  
 Summary statistics

ABSTRACT

In GWAS studies, SNP heritability measures the proportion of phenotypic variance explained by all measured SNPs. Accurate estimation of SNP heritability can help us better understand the degree to which measured genetic variants influence phenotypes. Over the last decade, a variety of statistical methods and software tools have been developed for SNP heritability estimation with different data types including genotype array data, imputed genotype data, whole-genome sequencing data, RNA sequencing data, and bisulfite sequencing data. However, a thorough technical review of these methods, especially from a statistical and computational viewpoint, is currently missing. To fill this knowledge gap, we present a comprehensive review on a broad category of recently developed and commonly used SNP heritability estimation methods. We focus on their modeling assumptions; their interconnected relationships; their applicability to quantitative, binary and count phenotypes; their use of individual level data versus summary statistics, as well as their utility for SNP heritability partitioning. We hope that this review will serve as a useful reference for both methodologists who develop heritability estimation methods and practitioners who perform heritability analysis.

© 2020 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

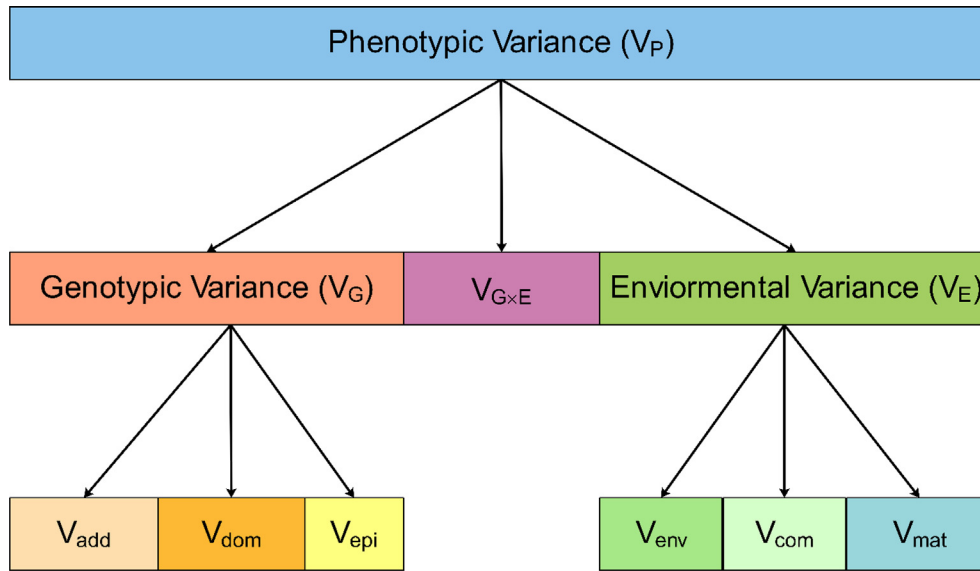
Contents

1. Introduction . . . . .	1557
2. Heritability estimation for quantitative traits . . . . .	1559
2.1. Notations . . . . .	1559
2.2. Modeling assumptions on the effect sizes . . . . .	1559
3. SNP heritability estimation for case control studies and count phenotypes . . . . .	1561
4. Inference algorithms for LMM and the adaptation of summary statistics . . . . .	1563
5. SNP heritability partitioning . . . . .	1565
6. Discussion . . . . .	1566
Declaration of Competing Interest . . . . .	1566
Acknowledgement . . . . .	1566
References . . . . .	1567

## 1. Introduction

A central quest in genetics is to understand the relative contribution of genetic factors and environmental factors to phenotypic

\* Corresponding author at: Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA.  
 E-mail address: [xzhousph@umich.edu](mailto:xzhousph@umich.edu) (X. Zhou).



**Fig. 1.** Partition of phenotypic variance.  $V_G$  represents the phenotypic variance due to genetic effects;  $V_E$  represents the phenotypic variance due to environmental effects; and  $V_{G \times E}$  represents the phenotypic variance due to gene-environment interactions. The genetic variance  $V_G$  can be partitioned into three parts:  $V_{add}$  that represents the additive genetic effects;  $V_{dom}$  represents the dominance genetic effects; and  $V_{epi}$  represents the epistatic effects. The environmental variance  $V_E$  can also be partitioned into three parts:  $V_{com}$  represents common environmental effects such as those due to residing in the same family;  $V_{mat}$  represents maternal effects such as nutritional intake during pregnancy; and  $V_{env}$  represents the residual stochastic environmental effects.

variation – a quest commonly framed as the nature vs nurture debate. Elucidating the relative contribution of genetics versus environment for various diseases and disease-related complex traits can help us better understand the causal mechanism of disease etiology and facilitate resource prioritization for disease diagnosis and prevention. A key quantity to evaluate the contribution of genetics versus environment is heritability, which measures the proportion of phenotypic variance explained by genetic factors. Two types of heritability can be estimated. The broad-sense heritability ( $H^2$ ) evaluates the proportion of phenotypic variance explained by all genetic factors, including additive effects, dominant effects, and epistasis effects. The narrow-sense heritability ( $h^2$ ), on the other hand, evaluates the proportion of phenotypic variance explained by additive genetic effects. Accurate estimation of heritability can show the degree to which genetic factors influence phenotypes and improve our understanding of the genetic basis of disease and disease-related complex traits. Indeed, heritability plays an important role across a range of genetic applications [1]: it is a key for understanding the evolutionary forces underlying natural selection; it determines how a population will respond to selection; it predicts, at least in part, gene mapping power in genome-wide association studies; it can estimate, quite accurately in some cases, the phenotypic value of an individual and thus facilitate genomic selection via predicted breeding values; and it provides an upper limit for the genetic prediction of phenotypes.

In the absence of genetic data, heritability estimation can be carried out in family/pedigree-based [2] or twin-based designs [3]. Recent reviews of heritability estimation in related individuals can be found from [4,5]. Briefly, in family/pedigree-based studies, heritability is often estimated using the linear mixed model (LMM), also known as the variance component model. The LMM allows partitioning of the phenotypic variance into different components: a genetic variance component, an environmental component, and potentially a gene-environment ( $G \times E$ ) interaction component (Fig. 1). The genetic variance component captures the part of phenotypic variance explained by genetic factors and can usually be further partitioned into three parts: the additive, dominance, and epistatic parts. The environmental variance component captures the part of phenotypic variance explained by environ-

mental factors and can also be further partitioned into three parts: the common environment part, the maternal influence part, and the residual stochastic environment part. The  $G \times E$  interaction variance component captures the part of phenotypic variance explained by the interactions between genetic factors and environmental factors. Measuring the  $G \times E$  component is often challenging as it requires the assessment of detailed environmental exposures obtained from appropriate study designs, with several recent analytic methods developed for modeling  $G \times E$  interactions for heritability estimation [6]. In this review, we focus on the genetic variance components.

After obtaining the estimates for various components, we can compute the estimated ratio of genetic variance over the total phenotypic variance, which serves as an estimate for heritability. For data with a relatively simple, familiar structure, a method of moments (MoM) based algorithm is often employed for variance component estimation in the linear mixed model. For example, a regression between the offspring phenotypic value ( $Y_O$ ) and the average of phenotypic value for the two parents ( $Y_P$ ) can directly lead to the heritability estimate ( $h^2 = 2 \cdot \text{Corr}(Y_O, Y_P)$ ) [7]. Similarly, for studies involving monozygotic (MZ) twins and dizygotic (DZ) twins, Falconer's formula is applied to obtain a heritability estimate, which equals twice the difference between the phenotypic correlation in MZ pairs and the phenotypic correlation in DZ pairs ( $H^2 = 2(r_{MZ} - r_{DZ})$ ) [8–10]. For family studies with individuals of different levels of relationship, a likelihood-based inference procedure is often employed for variance component estimation in the linear mixed model. This procedure constructs a kinship matrix based on pedigree information and then obtains the maximum likelihood estimate (MLE) or the restricted maximum likelihood estimate (REML) for heritability. Regardless of the estimation algorithm, various early family-based studies have produced accurate heritability estimates for various quantitative traits. For example, it has been estimated that the narrow-sense heritability for height is around 80% in family-based studies [1,11]. Heritability may vary across populations, across individual groups with different age or other characteristics, and may change over time. Similarly, heritability of livestock traits (e.g., milk yield) can sometimes double over a couple of decades through animal breeding programs [1].

The progress of array-based techniques, and more recently, whole-genome sequencing (WGS) techniques, have enabled accurate measurement of genotypes on millions of single nucleotide polymorphisms (SNPs) across the entire genome. These advances have subsequently enabled large-scale genome-wide association studies (GWASs) in apparently unrelated individuals. In the past decade, GWASs have identified many SNPs associated with various diseases and disease-related complex traits. However, the majority of identified SNPs only explain a small fraction of heritability for most traits, leading to a large fraction of unexplained heritability, commonly referred to “missing heritability”. Many explanations have been proposed for missing heritability. For example, it is possible the causal variants are not in complete linkage disequilibrium (LD) with the genotyped SNPs [12]. It is also possible that rare variants with large effects contribute disproportionately to the phenotypic variance. In addition, pedigree-based studies may have overestimated heritability [13]. A prominent hypothesis suggests that current GWASs are underpowered and that many causal SNPs remain undetected below the stringent genome-wide significant threshold, which can vastly underestimate the phenotypic variance. Therefore, it becomes critically important to estimate the heritability or phenotypic variance explained by all measured genome-wide SNPs. For example, in the seminal paper by Yang et al. [12], the heritability for height explained by significantly associated SNPs is only 10%, while that explained by all measured SNPs is near 50%, with an increase of 40%.

Several reviews have been recently written on SNP heritability estimation [4–5,14–20]. Most of these papers focus on the practical interpretation of SNP heritability and how it contributes to our understanding of missing heritability for various diseases and complex traits. In addition, these reviews are often targeted at experimental biologists, focus on quantitative traits, and have a narrow focus on a few existing statistical methods used for heritability estimation. To fill this critical gap, we provide a systematic review of various statistical methods that have been developed and used for SNP heritability estimation on quantitative traits, binary traits, and count phenotypes. Specifically, we included in our review all existing statistical methods that are developed for SNP heritability estimation. We did not include methods for heritability estimation in family/pedigree-based or twin-based designs. All methods in the review make use of genome-wide genotype data. Specifically, we focus on explaining the detailed modeling assumptions underlying various models, providing intuitions on how one would expect different models to work for traits with different genetic architectures. We include two commonly used algorithms -- the method of moments (MoM) and the restricted maximum likelihood (REML) -- the first of which allows estimation of SNP heritability using summary statistics. In addition, we include a broad category of methods that are suitable for modeling phenotypes in the form of traditional quantitative traits, ascertained case control status, and count measurements from recent genomic sequencing studies. We also include various methods for partitioning heritability across different SNP functional categories. We hope that our review can serve as a useful reference to the broad statistical genetics and computational biology communities on modeling and estimation of SNP heritability.

## 2. Heritability estimation for quantitative traits

### 2.1. Notations

First, we provide detailed notations and formulate the SNP heritability estimation problem into a statistical framework. We denote  $\mathbf{y} = (y_1, \dots, y_n)^T$  as the  $n$ -vector of quantitative trait measured on  $n$  individuals. We denote  $\mathbf{X}$  as the  $n$  by  $p$  matrix of

genotypes for the same  $n$  individuals and  $p$  SNPs. These genotypes can be collected in different forms, including genotype array data, imputed genotype data, and whole genome sequencing data. The genotype for the  $i$ -th individual and  $j$ -th SNP,  $\mathbf{X}_{ij}$ , is often coded as 0, 1, or 2, representing the number of reference alleles. For the genotype matrix, we assume that all missing genotypes have already been imputed with proper genotype imputation software; thus,  $\mathbf{X}_{ij}$  will be in the range of 0 and 2. To simplify the algebra, we further assume that each column of  $\mathbf{X}$  (i.e., SNP) has been centered to have zero mean; the results will remain the same with or without centering [21]. It is also common to standardize the columns of  $\mathbf{X}$  to have unit variance, which corresponds to making an assumption that rarer variants tend to have larger effects than common variants and variant effect sizes tend to decay with the inverse of the genotype variance [22]. Therefore, standardizing the columns will affect the results, although previous studies have shown that its relative contribution to the genetic relatedness matrix (GRM, defined later) is the same and has minimal influence on the SNP heritability estimation [23]. Therefore, we also standardize each column of  $\mathbf{X}$  to have unit variance. We use the following linear regression to model the relationship between the phenotype vector  $\mathbf{y}$  and the genotype matrix  $\mathbf{X}$ ,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (1)$$

where  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$  is an  $p$ -vector of genetic effect sizes and  $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^T$  is the  $n$ -vector of residual errors with each entry  $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ . Note that centering of the phenotype  $\mathbf{y}$  and each column of genotype matrix  $\mathbf{X}$  allows us to ignore the intercept in equation (1).

Equation (1) is often used for estimating the proportion of phenotypic variance explained by all measured SNPs in GWAS,  $\text{Var}(\mathbf{X}\boldsymbol{\beta})/\text{Var}(\mathbf{y})$ , where  $\text{Var}$  denotes sample variance. This quantity is commonly referred to as the proportion of variance in phenotypes explained (PVE) by available genotypes or SNP heritability, denoted as  $h_g^2$ . A simple approach to estimate SNP heritability is to select a candidate set of associated SNPs and then estimate PVE by these selected SNPs. For example, for a given trait, we can identify SNPs that are associated with the trait passing the genome-wide p-value significance threshold of  $5 \times 10^{-8}$ . We can include all these significant SNPs, or further extract uncorrelated SNPs from the set through linkage disequilibrium (LD) pruning or clumping, into the model defined in equation (1). When the number of SNPs,  $p$ , is small, we can easily estimate  $\boldsymbol{\beta}$  through ordinary least square estimation. When the selected SNPs are independent from each other, we can estimate the PVE as  $\widehat{h_g^2} = \sum_{j \in A} \widehat{\beta}_j^2$ , where  $A$  denotes the set of selected independent SNPs; when SNPs are correlated with each other, we can estimate the PVE using the definition  $\text{Var}(\mathbf{X}\boldsymbol{\beta})/\text{Var}(\mathbf{y})$ . As noted above, such estimation may greatly underestimate the true SNP heritability because it ignores many SNPs that have not reached the stringent genome-wide threshold due to imperfect statistical power. For example, the top ~ 50 SNPs only explain 5–10% of phenotypic variance for height [24–27] although, based on pedigree-based designs, height is a highly heritable trait with estimated heritability as high as 80% [1,11].

### 2.2. Modeling assumptions on the effect sizes

When we include all genome-wide SNPs, the number of SNPs  $p$  would exceed the number of individuals  $n$ . Indeed, in a typical GWAS,  $p$  is often on the order of a million to 100 million while  $n$  typically ranges from a few thousand to half a million. When  $p \gg n$ , the regression model defined in equation (1) becomes an underdetermined system. Therefore, we need to make certain

modeling assumptions on the SNP effect sizes  $\beta$  to complete the modeling specification. Various modeling assumptions have been proposed on the effect sizes  $\beta$ . We describe a few commonly used modeling assumptions below.

**Bayesian variable selection regression (BVSR).** Perhaps the first attempt in the genetics field to estimate SNP heritability was BVSR. BVSR makes a sparse modeling assumption that a relatively small proportion of all genetic variants truly affect the phenotype [28–33]. In particular, BVSR [32] assumes that the genetic effect size of each SNP follows a point-normal distribution

$$\beta_j \sim \pi N(0, \sigma_\beta^2) + (1 - \pi)\delta_0, \quad (2)$$

where  $\sigma_\beta^2$  is the variance component in the normal distribution and  $\delta_0$  denotes a point mass at zero. The point normal is commonly referred to as a spike and slab prior. Based on the point-normal modeling assumption, with proportion  $\pi$ , the SNP effect size is non-zero and follows a normal distribution; while with proportion  $1 - \pi$ , the SNP effect size is exactly zero. The proportion of non-zero effect SNPs,  $\pi$ , is often assumed to be small. For example, BVSR [32] places a sparsity-inducing prior on  $\pi$  and assumes that  $\log(\pi)$  follows a uniform distribution *a priori*. The model defines a parameter to quantify PVE,  $h_g^2 = \frac{p^* \sigma_\beta^2}{p^* \sigma_\beta^2 + \sigma_e^2}$ , with  $p^*$  as the number of non-zero  $\beta$ .

Given the observed data, BVSR relies on a Markov Chain Monte Carlo (MCMC) algorithm to obtain posterior samples from the approximate posterior distribution for SNP heritability. As an example, BVSR was applied to analyze a GWAS study with plasma C-reactive protein (CRP) and found that using all SNPs explained approximately 6% of variance in CRP, much higher than the previous estimates of 0.35%–0.4% using only significant SNPs [34–35].

**Linear mixed model (LMM).** The second modeling assumption, which is most often used, is the normality assumption on the effect sizes. This modeling assumption is commonly referred to as LMM. LMM was first proposed by [12] for SNP heritability estimation, after which LMM became a standard method and one of the most effective approaches for the analytic task. LMM assumes that

$$\beta_j \sim N(0, \sigma_\beta^2) \quad (3)$$

Under this modeling assumption, all SNPs have non-zero effects with their effect sizes following a normal distribution. LMM is commonly used for heritability estimation as well as association analysis while accounting for family relatedness or population stratification. When a particular likelihood-based inference procedure REML (more details below) is used to perform parameter inference, LMM is also known as the ridge regression or Genome-based REML (GREML). In this review, the notations of LMM, GREML, REML, and GCTA [36] all represent this approach. In the LMM, the variance of total additive genetic effects can be defined as  $\sigma_g^2 = p \cdot \sigma_\beta^2$ . The variance components  $\sigma_g^2$  and  $\sigma_e^2$  can be estimated using software such as GCTA [36] and GEMMA [37]. The SNP heritability estimate can be expressed as  $\widehat{h}_g^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2}$ . Note that, when the columns of the genotype matrix  $\mathbf{X}$  are not standardized, a scaling factor  $s = \text{trace}(\mathbf{K})/n$ , where  $\mathbf{K}$  denotes the estimated genetic relatedness matrix (GRM) and is commonly computed as  $\mathbf{K} = \frac{1}{p} \mathbf{X} \mathbf{X}^T$ , is multiplied by  $\sigma_g^2$ , which leads to the SNP heritability estimate as  $\widehat{h}_g^2 = \frac{s \sigma_g^2}{s \sigma_g^2 + \sigma_e^2}$  [22].

**Bayesian sparse linear mixed model (BSLMM).** Because BVSR and LMM make completely different modeling assumptions, one may naturally expect that the two models work better for traits with different genetic architectures. Specifically, because of the sparse effect size assumption, BVSR is more accurate in estimating SNP heritability when a small proportion of SNPs truly has non-zero

effects on the trait. In contrast, because of the polygenic effect size assumption, LMM is more accurate in estimating SNP heritability when truly a large proportion of SNPs have non-zero effects on the trait. Thus, BVSR tends to underestimate SNP heritability for polygenic traits while LMM tends to be imprecise for non-polygenic traits [22] – even though LMM produces unbiased estimates for traits with various genetic architectures [22,38]. Unfortunately, the true genetic architecture of a phenotype is unknown *a priori*. Therefore, it is often unclear whether one should use LMM or BVSR to analyze a given trait.

Motivated by this methodological limitation, Zhou et al. [22] proposed a hybrid model of LMM and BVSR, which is referred to as Bayesian sparse linear mixed model (BSLMM). BSLMM places a mixture of two normal distributions on the effect sizes,

$$\beta_j \sim \pi N(0, \sigma_a^2 + \sigma_b^2) + (1 - \pi)N(0, \sigma_b^2). \quad (4)$$

That is, with probability  $1 - \pi$ ,  $\beta_j$  tends to be small and follows a normal distribution with a small background variance of  $\sigma_b^2$ ; while with probability  $\pi$ ,  $\beta_j$  tends to be large and follows a normal distribution with a large variance of  $\sigma_a^2 + \sigma_b^2$ , where  $\sigma_a^2$  is the additional variance on top of the background variance. Clearly, when  $\pi = 0$ , BSLMM reduces to LMM. When  $\sigma_b^2 = 0$ , BSLMM reduces to BVSR. By including both LMM and BVSR as special cases, BSLMM can take advantage of LMM and BVSR to adaptively infer the genetic architecture underlying the trait from the data at hand. In the BSLMM, the SNP heritability can be expressed as a population level parameter  $h_g^2 = \frac{p\pi\sigma_a^2 + \sigma_b^2}{p\pi\sigma_a^2 + \sigma_b^2 + \sigma_e^2}$ , the approximate expectation of PVE. Besides

$h_g^2$ , BSLMM also defines a parameter  $\rho = \frac{p\pi\sigma_a^2}{p\pi\sigma_a^2 + \sigma_b^2}$  as the approximate expectation of PGE, the proportion of genetic variance explained by the sparse effects. In addition, PVE can be estimated by  $\frac{\text{Var}(\mathbf{X}\beta + \mathbf{u})}{\text{Var}(\mathbf{y})}$

while PGE can be estimated by  $\frac{\text{Var}(\mathbf{X}\beta)}{\text{Var}(\mathbf{X}\beta + \mathbf{u})}$ , where  $\mathbf{u}$  is the random effect following  $\mathbf{u} \sim N(\mathbf{0}, \sigma_b^2 \mathbf{K})$ . BSLMM relies on a Metropolis Hastings (MH) algorithm to perform posterior inference. BSLMM is also closely related to the recent omnigenic model hypothesis [39]. Specifically, the omnigenic model hypothesizes that all genes have non-zero effects, which is modeled in BSLMM by assuming that all SNPs have non-zero effects. In addition, the omnigenic model hypothesizes that a small proportion of genes, denoted as core genes, have additional effects. These additional effects are modeled by the normal component with a large variance in BSLMM. As an extension of BSLMM, Zhu and Stephens [40] provided a summary statistics-based version, Regression with Summary Statistics (RSS) likelihood. RSS likelihood allows BSLMM to be applied to large scale GWASs. By analyzing a summary-level GWAS with 253,288 individuals genotyped at 1.06 million SNPs using BSLMM, RSS likelihood obtained the heritability estimate for height as 52% [40].

**Linkage disequilibrium adjusted kinships (LDAK).** The above BVSR, LMM/GREML and BSLMM assume that the effect size for the  $j$ -th SNP,  $\beta_j$ , does not depend on how many SNPs are in close linkage disequilibrium (LD) with the  $j$ -th SNP. In contrast, the Linkage Disequilibrium Adjusted Kinships (LDAK) [41–42] assumes that  $\beta_j$  depends on how  $j$ -th SNP is correlated with its neighborhood SNPs. Specifically, similar to LMM, LDAK assumes that  $\beta_j$  follows a normal distribution  $\beta_j \sim N(0, \sigma_j^2)$ . However, different from LMM that assumes the same variance  $\sigma_j^2 = \sigma^2$ , LDAK assumes that  $\sigma_j^2$  is  $j$ -th SNP specific and is related to minor allele frequency, LD score, and imputation information score of the SNP. The LD score of SNP  $j$  is defined as  $\ell_j = \sum_{j=1}^p r_{jj}^2 - \frac{1}{n}$ , where  $\sum_{j=1}^p r_{jj}^2$  is the sum of the squared Pearson's correlation between SNP  $j$  and all other SNP's  $j$  while  $\frac{1}{n}$  represents the expectation of the summation under the



null and a high value indicates that the  $j$ -th SNP is in high LD with many nearby SNPs. The imputation information score is a metric between 0 and 1 output from imputation software: a value of 1 indicates that there is no uncertainty in the imputed genotypes while a value of 0 means that there is complete uncertainty about the genotypes. Specifically, LDAK assumes that the effect size  $\beta_j$  follows

$$\beta_j \sim N(0, \sigma_j^2), \sigma_j^2 \propto [f_j(1 - f_j)]^{1+\alpha} \times w_j \times r_j, \quad (5)$$

where  $f_j$  is the minor allele frequency of the SNP  $j$ ;  $w_j$  is a SNP-specific weight that is a function of the inverse of the LD score of SNP  $j$ , so that the  $j$ -th SNP effect size tends to be smaller if there are more SNPs in LD with the  $j$ -th SNP; and  $r_j \in [0, 1]$  is the imputation information score measuring genotype certainty, so that the  $j$ -th SNP effect size tends to be smaller for the genotype with higher uncertainty. The parameter  $\alpha$  determines the relationship between  $\sigma_j^2$  and  $f_j$ . Specifically,  $\alpha = -1$  indicates that  $\sigma_j^2$  does not depend on  $f_j$ , an assumption commonly made in genetics;  $\alpha < -1$  (e.g.,  $-1.25$ ) indicates that  $\sigma_j^2$  decreases as  $f_j$  increases; and  $\alpha > -1$  (e.g.,  $-0.75, -0.25$ ) indicates that  $\sigma_j^2$  increases as  $f_j$  increases. The default value of  $\alpha$  in LDAK is  $-0.25$ . LDAK relies on REML to estimate parameters. Because of different modeling assumptions of LDAK and LMM, different SNP heritability estimations are obtained by different methods in real data analysis. For example, if the underlying SNP effect size depends on LD in the same form as of LDAK, then methods, such as LMM that fails to model the effect size on LD score dependency, would generate downward biased estimates. Indeed, in a real data application, LDAK obtained an average of 43% SNP heritability estimation higher than that of LMM for 19 analyzed traits [42]. Certainly, while the naïve LMM does not account for the potential LD dependency, it also can be extended to do so by LD stratified analysis; such extensions are described in the SNP Heritability Partitioning section.

Besides these above methods, several other models can be used for SNP heritability estimation. Particularly, many phenotype prediction models developed elsewhere can be directly applied for SNP heritability estimation. For example, the Bayesian alphabet models assume that the genetic effect sizes follow either a  $t$ -distribution (BayesA) [43–46], a mixture of  $t$ -distribution and a point mass at zero (BayesB, BayesD, BayesD $\pi$ ) [43–47], or a mixture of two  $t$ -distributions (BayesC) [43,45]. The Bayesian lasso assumes a double exponential distribution [28,48]. BayesR assumes a mixture of three normal distributions and a point mass at zero [49]. NEG assumes a normal exponential gamma distribution [29]. BayesS assumes a point-normal distribution with SNP-specific variance as a function of MAF [50].

*Dirichlet process regression (DPR).* These aforementioned methods share a common feature of relying on a finite number of parameters to characterize the genetic effect distribution; that is, they all use parametric models. In contrast to the parametric model, Zeng and Zhou [51] developed a Bayesian non-parametric model, termed as the latent Dirichlet process regression (DPR). DPR assumes that  $\beta_j$  follows a normal distribution, with a further unknown distribution  $G$  placed upon the variance parameter  $\sigma_j^2$ . DPR actively infers the unknown distribution  $G$  by assuming a non-parametric Dirichlet process (DP) prior on the distribution itself:

$$\beta_j \sim N(0, \sigma_j^2), \sigma_j^2 \sim G, G \sim DP(IG(a, b), \lambda) \quad (6)$$

where the inverse gamma (IG) distribution is the base distribution and the concentration parameter  $\lambda$  determines how the distribution of  $G$  differs from the base distribution. By inferring the

distribution  $G$  based on the data at hand, DPR is flexible and adaptive to a wide range of genetic architectures, resulting in accurate phenotype prediction and appreciable power gain for the transcriptome-wide association studies (TWAS) [52–53], an integrative analysis of expression mapping studies and GWASs. Note that the above modeling assumption is also equivalent to assuming each element of  $\beta$  follows a mixture of infinitely many normal distributions *a priori*,

$$\beta_j \sim \sum_{\varphi=1}^{+\infty} \pi_{\varphi} N(0, \sigma_{\varphi}^2), \pi_{\varphi} = v_{\varphi} \prod_{l=1}^{\varphi-1} (1 - v_l), v_{\varphi} \sim \text{Beta}(1, \lambda) \quad (7)$$

Here,  $\pi_{\varphi}$  is the weight corresponding to the  $\varphi$ -th normal distribution; it is generated from a stick breaking process and determined by a latent proportion parameter  $v_l$  that each follows a Beta prior. With the DPR modeling assumption, one can obtain the SNP heritability estimates via two algorithms: the Monte Carlo Markov Chain and the variational Bayesian algorithm.

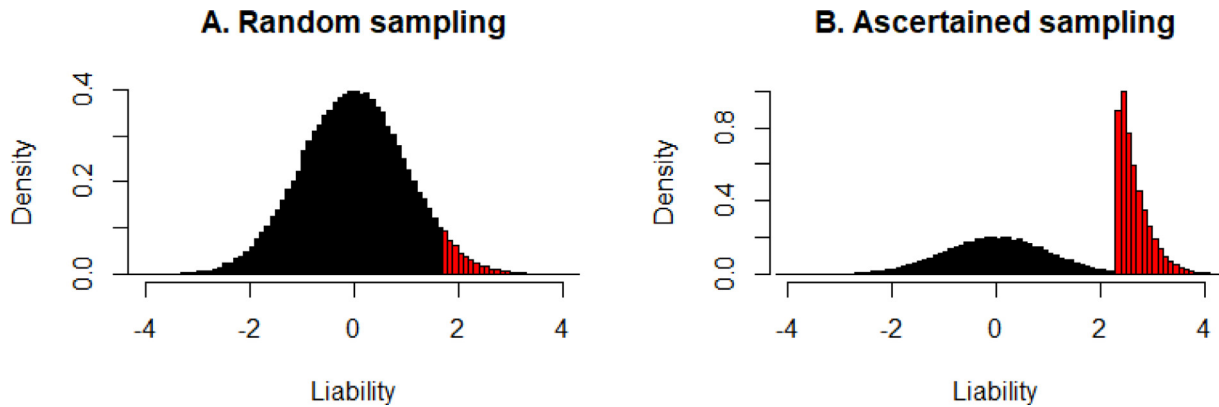
### 3. SNP heritability estimation for case control studies and count phenotypes

*Liability threshold model: REML.* We have focused on estimation of SNP heritability for quantitative traits. More considerations are needed when the outcome is a disease phenotype obtained from case control studies. In this case, estimation of SNP heritability requires not only proper modeling of the binary nature of the outcome, but also proper controlling of the ascertainment occurred in case control studies. The binary nature of case control status suggests that the variance of the phenotype is a function of its mean, rendering invalid normality assumption on the residual errors. The normality assumption is commonly used in SNP heritability estimation for quantitative traits as described in previous sections. Ascertainment is a result of the case-control sampling design where the proportion of cases in the study is collected to be much higher than that in the population. Ascertainment effectively increases the associated SNP effect size estimates as compared to that in the population; it renders inaccurate effect size assumptions made in SNP heritability estimation for quantitative traits. Therefore, methods for quantitative traits are no longer applicable to case control studies. Instead, a liability threshold model is used to account for both the binary nature and ascertainment in case control studies.

The liability threshold model was first introduced by [54]. It introduces a latent continuous variable for every individual  $i$ , termed as liability score  $l_i$ . The liability score effectively measures the individual's susceptibility to disease. The liability score, paired with the liability threshold value of  $\ell$ , determines whether an individual is a case or a control: the individual  $i$  is a case when  $l_i > \ell$  and is a control otherwise. The liability score is a continuous variable and assumed to follow the same linear model as in equation (1)

$$l = X\beta + \varepsilon, \quad (8)$$

where  $X\beta$  represents the genetic contribution to liability and  $\varepsilon$  represents the environmental contribution to liability. As described in previous sections, different modeling assumptions can be made on the genetic effect sizes  $\beta$ , though the common choice is the normal assumption. Under this assumption, the liability score follows a (multivariate) normal distribution in the population (Fig. 2A). The liability threshold  $\ell$ , when paired with a distributional assumption on  $l_i$ , effectively determines the prevalence of the disease. Certainly, due to ascertainment, the liability scores in the case control study no longer follow a normal distribution but are often enriched with large liability scores (Fig. 2B). To



**Fig. 2.** The distributions of liability score under random sampling and ascertained case-control sampling for unrelated individuals. Red represents cases while black represents controls. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

address the issue of nonnormality with liability score, Lee et al. [55] proposed a transformation procedure. In particular, the transformation procedure first treats the disease status (0/1) as a continuous outcome and fits an LMM described in equations (1) and (3) using the REML algorithm to estimate the variance components. The resulting SNP heritability estimate  $h_o^2$  is referred to as the SNP heritability estimated on the observed scale. Afterwards, the transformation procedure applies a linear transformation on  $h_o^2$  and converts it to the SNP heritability estimate  $h_g^2$  on the liability scale:

$$\widehat{h}_g^2 = \frac{P^2(1-P)^2}{P^*(1-P^*)\varphi(\ell)^2} \widehat{h}_o^2 \quad (9)$$

where  $\varphi(\cdot)$  is the standard normal density function,  $\widehat{h}_o^2$  is the REML heritability estimates obtained by treating the case-control status (0/1) as a quantitative trait,  $P$  represents the disease proportion in the population, and  $P^*$  denotes the disease proportion in the sample. The above transformation extends the Dempster and Lerner [56] formula  $h_g^2 = \frac{P(1-P)}{\varphi(\ell)^2} \cdot h_o^2$  which addresses the binary nature of case control outcome but does not account for ascertainment.

**Liability threshold model: HE/PCGC.** The detailed algorithmic derivation in Lee et al. [55] is complicated. Based on Taylor series expansion and approximation, Zhou et al. [22] provided an alternative derivation, which led to the same transformation equation described in equation [9]. However, the new derivation casts concern on the effectiveness of such transformation when paired with REML, as the approximation in equation [9] is only valid when SNP heritability is close to zero. When SNP heritability is not close to zero, the SNP heritability on the liability scale based on equation [9] will be underestimated [57–58]. Golan et al. [58] provides a simple solution to the downward bias in SNP heritability estimation: instead of using REML estimates, one can use the Haseman-Elston (HE) regression to obtain the variance component estimates. The HE regression is also referred to as the phenotype correlation-genotype correlation (PCGC) regression relying on the equation  $E(P_{i,j}) = f(h_g^2, G_{i,j})$ , where  $P_{i,j}$  is the phenotypic correlation between individual  $i$  and  $j$  and  $G_{i,j}$  is the genotypic correlation between the two individuals. Here,  $f(\cdot)$  is a function that relates SNP heritability and genetic correlation  $G_{i,j}$  to the phenotypic correlation  $P_{i,j}$ . The specific functional form of  $f(\cdot)$  depends on the design of the study and the properties of the phenotype. In HE regression, the function  $f(\cdot)$  is a simple product of the SNP heritability and genetic correlation; that is,  $f(h_g^2, G_{i,j}) = h_g^2 \cdot G_{i,j}$ . Instead of requiring low SNP

heritability, HE/PCGC regression only requires each element in the kinship matrix close to zero. Consequently, it can be widely applied to data collected on unrelated individuals and provides approximately unbiased SNP heritability estimates on the liability scale in several GWASs [58]. The HE/PCGC regression is later recognized to be linked to the MINQUE (the minimal norm quadratic unbiased estimation) estimation proposed in the statistical literature and can be viewed from a method of moments (MoM) algorithm perspective [59]. The HE/PCGC regression is further extended to model ascertained case control studies using summary statistics [59–60].

**Generalized linear mixed model: PQLseq.** Besides binary traits and case control studies, many complex traits are measured on various other data types through genomic sequencing studies. For example, RNA sequencing (RNAseq) studies have allowed accurate gene expression measurements across tens of thousands of genes. Bisulfite sequencing (BSseq) studies have enabled accurate methylation profiling across genome wide CpG sites. Understanding SNP heritability of these molecular traits, including gene expression levels and methylation levels, can facilitate our understanding of the causal or mediation mechanism underlying the SNP-trait associations. These two types of sequencing data have different data structures. Specifically, RNAseq studies collect one read count for each gene as its expression level. In contrast, BSseq studies collect two read counts for each CpG site – one methylated count and one total count – as the methylation level. The ratio between these two counts represents approximately the methylation proportion of the given CpG site. Both types of data are of count nature. The standard SNP heritability estimation method, LMM, has been recently applied to estimate heritability of gene expression [61–65], of methylation level [66–68], and of various other molecular traits [69]. However, LMM is specifically designed for analyzing quantitative traits. In genomic sequencing studies, the application of LMM requires *a priori* transformation of the count data to continuous data before heritability estimation [61,70]. Transforming sequencing count data may fail to account for the sampling noise from the underlying count generating process and may inappropriately attribute such noise to independent environmental variation. As shown in Sun et al. [71], modeling count data with LMM can run into the risk of overestimating environmental variance and subsequently underestimating heritability. To mitigate the problem, Sun et al. [71] developed PQLseq, a penalized quasi-likelihood for sequencing count data, based on the generalized linear mixed models (GLMM), which directly model count data. For a given gene in an RNAseq study, PQLseq considers a Poisson mixed model (PMM) to directly model the count data  $y_i \sim Poi(N_i \lambda_i)$  for the  $i$ -th individual, where  $y_i$  is the number of reads mapped to the partic-

ular gene,  $N_i$  is the total read counts (a.k.a read depth or coverage), and  $\lambda_i$  is an unknown Poisson rate parameter that represents the underlying gene expression level. For a given CpG site in a BSseq study, PQLseq considers a binomial mixed model (BMM)  $y_i \sim \text{Bin}(r_i, \pi_i)$ , where  $r_i$  is the total read count for the  $i$ -th individual,  $y_i$  is the methylated read count constrained to be an integer value less than or equal to  $r_i$ , and  $\pi_i$  is an unknown parameter that represents the underlying proportion of methylated reads at the site. For either model, PQLseq transforms the unknown parameters into a latent variable  $z_i$ :  $z_i = \log(\lambda_i)$  in PMM and  $z_i = \text{logit}(\pi_i)$  in BMM. The latent variable  $z_i$  is combined together into a vector, which is modelled as follows:

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (10)$$

where  $\mathbf{X}\boldsymbol{\beta}$  represents the genetic contribution to the latent variable  $\mathbf{z}$  and  $\boldsymbol{\varepsilon}$  represents the environmental contribution. PQLseq further relies on the penalized quasi-likelihood for parameter inference to obtain unbiased SNP heritability estimates.

#### 4. Inference algorithms for LMM and the adaptation of summary statistics

All the methods described so far require individual-level genotypes and phenotypes data from all samples in the study. Because of consent and privacy concerns, and logistic limitations (e.g., large-scale data transfer and storage often require high-end computing infrastructure), it is increasingly difficult to access complete individual-level data from large-scale association studies. Indeed, sharing summary statistics such as the marginal z-scores across multiple studies, performing meta-analysis, and releasing results in terms of summary statistics has become a standard practice in most consortium studies. Requiring complete individual-level data restricts the use of many SNP heritability estimation methods and limits their benefits in many large-scale studies. In addition, the aforementioned methods are computationally expensive. For example, the REML algorithm in LMM or GLMM scales cubically with respect to the sample size. Similarly, both BVS and BSLMM require computationally expensive Markov chain Monte Carlo methods for model fitting. To alleviate the computational concern and make use of summary statistics, several alternative statistical methods for SNP heritability have been recently developed.

A common method to estimate SNP heritability based on summary-statistics is LD Score regression (LDSC) [72]. For each SNP, LDSC first computes its LD score,  $\ell_j$ , which is defined in the above LDK section and captures approximately the number of genetic variants tagged by this SNP. LD score cannot be computed exactly due to the large number of genome-wide SNPs. Instead, it is typically estimated based on SNPs in an appropriate sliding window (e.g., 1 MB or 1 cM). After obtaining LD score, LDSC regresses the  $\chi^2$  test statistic from GWAS on the per-SNP LD scores

$$E[\chi_j^2 | \ell_j] = n\ell_j \cdot \frac{h_g^2}{p} + na + 1, \quad (11)$$

where  $a$  measures the confounding bias due to potential population stratification and cryptic relatedness. Here, population stratification refers to the presence of a systematic difference in allele frequencies between subpopulations in the data possibly due to different ancestry. Cryptic relatedness occurs when individuals in the study are more closely related to another than thought. Both population stratification and cryptic relatedness, if uncontrolled, can lead to upward biased SNP heritability estimation. By controlling for population stratification and cryptic relatedness using the parameter  $a$ , LDSC can mitigate their influence for SNP heritability estimation. Thus, regressing the GWAS test statistics  $\chi_j^2$  on per-SNP LD scores  $\ell_j$  allows for estimation of  $h_g^2$ . Unlike standard data-generative models

(i.e., models that describe how the individual-level variables  $\mathbf{y}$  are generated based on genotypes of  $n$  samples), LDSC models the marginal test statistics for  $p$  SNPs. By modeling summary statistics, LDSC is not only applied to many data sets that previously cannot be analyzed for SNP heritability estimation, it also substantially improves computational speed and makes SNP heritability scalable to large data sets. LDSC was initially introduced without an underlying data-generative model. It was later found out that LDSC is fitting the LMM described in [59]. However, instead of applying the standard likelihood-based approach REML for fitting LMM, LDSC relies on a matching moments-based method. From this aspect, LDSC is closely related to HE/PCGC methods.

Zhou [59] developed MQS (MinQue for Summary statistics) and related it with LDSC and HE/PCGC. MQS is based on the MINQUE criterion, a conceptual framework based on MoM. For the case of one variance component, an analytic variance component estimation is as follows:

$$\hat{\sigma}_\beta^2 = S^{-1}q, \quad (12)$$

$$\text{where } q = \frac{\mathbf{y}^T(\mathbf{A} - \mathbf{I})\mathbf{y}}{(n-1)^2}, S = \frac{\text{tr}(\mathbf{A}\mathbf{K})}{(n-1)^2} - \frac{1}{n-1} \quad (13)$$

where  $\mathbf{A}$  is an  $n$  by  $n$  matrix. All choices of  $\mathbf{A}$  can lead to unbiased variance component estimates while different choices of  $\mathbf{A}$  can influence the estimation accuracy. In particular, the optimal choice  $\mathbf{A} = (\sigma_\beta^2 \mathbf{K} + \mathbf{I})^{-1} \mathbf{K} (\sigma_\beta^2 \mathbf{K} + \mathbf{I})^{-1}$  with known  $\sigma_\beta^2$  leads to the most accurate variance component estimates. In practice,  $\sigma_\beta^2$  is unknown and the optimal choice of  $\mathbf{A}$  cannot be used. Therefore, we will need to make decisions on the choice of  $\mathbf{A}$ . Different choices of  $\mathbf{A}$  in MQS lead to different existing variance component estimation algorithms. Specifically, when  $\mathbf{A} = (\hat{\sigma}_\beta^2 \mathbf{K} + \mathbf{I})^{-1} \mathbf{K} (\hat{\sigma}_\beta^2 \mathbf{K} + \mathbf{I})^{-1}$  and  $\hat{\sigma}_\beta^2$  is updated through the above estimation equation in an iterative fashion, MQS becomes REML. When  $\mathbf{A} = \mathbf{X}\mathbf{W}\mathbf{X}^T/p$  with a certain diagonal weighting matrix  $\mathbf{W}$ , MQS becomes the weighted version of LDSC. When  $\mathbf{A} = \mathbf{K}$ , MQS becomes HE/PCGC. MQS brings many seemingly unrelated methods – REML, HE/PCGC, LDSC – into the same unified statistical framework. With this new framework, MQS provides an alternative but mathematically equivalent form of HE/PCGC to allow for the use of summary statistics. MQS also provides an exact approximation of LDSC for yielding unbiased and statistically more efficient SNP heritability estimation. In addition, MQS can be easily extended to model multiple variance components or multiple phenotypes. Finally, while MQS requires computing  $q$  in equation [12] using all individuals, it can use only a subset of individuals to estimate  $S$  without incurring accuracy loss for the final SNP heritability estimates. Such strategy of MQS, using a subset of individuals for estimating certain quantities while using all individuals for computing other quantities, is in line with the idea of stochastic approximation as in Robbins and Monro [73]. The stochastic estimation strategy used in MQS leads to computational speed improvement over standard methods by orders of magnitude.

While both LDSC and MQS rely on the standard LMM assumption, the recently proposed SumHer [74] makes a different modeling assumption on the SNP effect sizes based on the LDK model. SumHer effectively extends LDK [41–42] to use summary statistics. Another method extended from existing approach is PCGC-s [60], which extends the PCGC approach [58] to use summary statistics as well as the genetic correlations between two diseases. A summary of methods for estimating SNP heritability is shown in Table 1 and a corresponding decision tree is in Fig. 3.

**Table 1**  
A summary of methods for SNP heritability estimation.

Main Text Sections	Methods	Modeling Assumptions	Estimation Algorithms	Trait Types	Software	Weblink	Comments	References
Modeling assumptions	BVSR	$\beta_j \sim \pi N(0, \sigma_\beta^2) + (1 - \pi)\delta_0$	MCMC	Quantitative	GEMMA	<a href="https://github.com/genetics-statistics/GEMMA">https://github.com/genetics-statistics/GEMMA</a>	Fast for large-scale data; Also useful for phenotype prediction and PRS construction; supports Mac and Linux platforms	[32–33]
	BSLMM	$\beta_j \sim \pi N(0, \sigma_a^2 + \sigma_b^2) + (1 - \pi)N(0, \sigma_b^2)$	MCMC				Also useful for SNP association tests with LMM; supports Windows, Mac and Linux platforms	[22]
	LMM/REML	$\beta_j \sim N(0, \sigma_\beta^2)$	REML	Quantitative	GEMMA/GCTA	<a href="https://github.com/genetics-statistics/GEMMA">https://github.com/genetics-statistics/GEMMA</a> <a href="http://cnsgenomics.com/software/gcta/#Overview">http://cnsgenomics.com/software/gcta/#Overview</a>		[12]
	LDAC	$\beta_j \sim N(0, \sigma_j^2), \sigma_j^2 \propto w_j (f_j(1 - f_j))^{0.75}$	REML	Quantitative	LDAC	<a href="http://dougsspeed.com/ldac/">http://dougsspeed.com/ldac/</a>	Over 30 functions, importantly, SNP heritability estimation and SNP-based prediction models construction, supported for Mac and Linux platforms	[41–42]
	DPR	$\beta_j \sim \sum_{\phi=1}^{+\infty} \pi_\phi N(0, \sigma_\phi^2), \sigma_\phi^2 \sim G, G \sim DP(H, \lambda)$	MCMC/VB	Quantitative	DPR	<a href="https://github.com/biostatpzeng/DPR">https://github.com/biostatpzeng/DPR</a>	Mainly for robust genetic prediction and PRS construction of complex traits; supports Mac and Linux platforms	[51]
Case-control study	HE/PCGC	$\beta_j \sim N(0, \sigma_j^2)$	MoM	Binary	PCGC	<a href="https://data.broadinstitute.org/alkesgroup/PCGC/">https://data.broadinstitute.org/alkesgroup/PCGC/</a>	Mitigate biases in REML heritability estimation for ascertained case-control studies; supports Linux platform	[58]
Count data	PQLseq	$\beta_j \sim N(0, \sigma_\beta^2)$	MCMC	Binary/Count	PQLseq	<a href="https://cran.r-project.org/web/packages/PQLseq/index.html">https://cran.r-project.org/web/packages/PQLseq/index.html</a>	For heritability estimation of count data in RNAseq and Bisulfite seq studies; supports Windows, Mac and Linux platforms	[71]
Summary statistics	LDSC	$E(\beta_j) = 0, \sigma_j^2 = h_g^2/p$	MoM	Quantitative/Binary	LDSC	<a href="https://github.com/bulik/ldsc">https://github.com/bulik/ldsc</a>	A command tool for estimating heritability and genetic correlation using GWAS summary statistics	[72]
	MQS	$\beta_j \sim N(0, \sigma_j^2)$	MoM	Quantitative/Binary	GEMMA	<a href="https://github.com/genetics-statistics/GEMMA">https://github.com/genetics-statistics/GEMMA</a>	A general statistical framework for SNP heritability estimation using summary statistics; supports Mac and Linux platforms	[59]
	SumHer	$E(\beta_j) = 0, \text{Var}(\beta_j) \propto w_j (f_j(1 - f_j))^{0.75}$	REML	Quantitative/Binary	SumHer	<a href="http://dougsspeed.com/sumher/">http://dougsspeed.com/sumher/</a>	Heritability estimation using summary statistics under the LDAC assumption; supports Linux platform	[74]

Table lists 10 methods described in the main text, with the first seven methods for analyzing individual level data and last three methods for analyzing summary statistics. Columns contain the main text section in which the method is described (1st column), method name (2nd column), modeling assumption on the SNP effect sizes (3rd column), estimation algorithms (4th column), phenotype type (5th column), implemented software (6th column), web link (7th column), additional comments (8th column) and references (9th column). In the 3rd column,  $\delta_0$  denotes a point mass at zero;  $N(\dots)$  denotes a normal distribution with the mean and variance parameters; DP denotes a Dirichlet process. In the 4th column, MCMC represents Markov chain Monte Carlo method, VB represents variational Bayesian, REML represents restricted maximum likelihood method, and MoM represents method of moments. In the 8th column, PRS is short for polygenic risk scores.



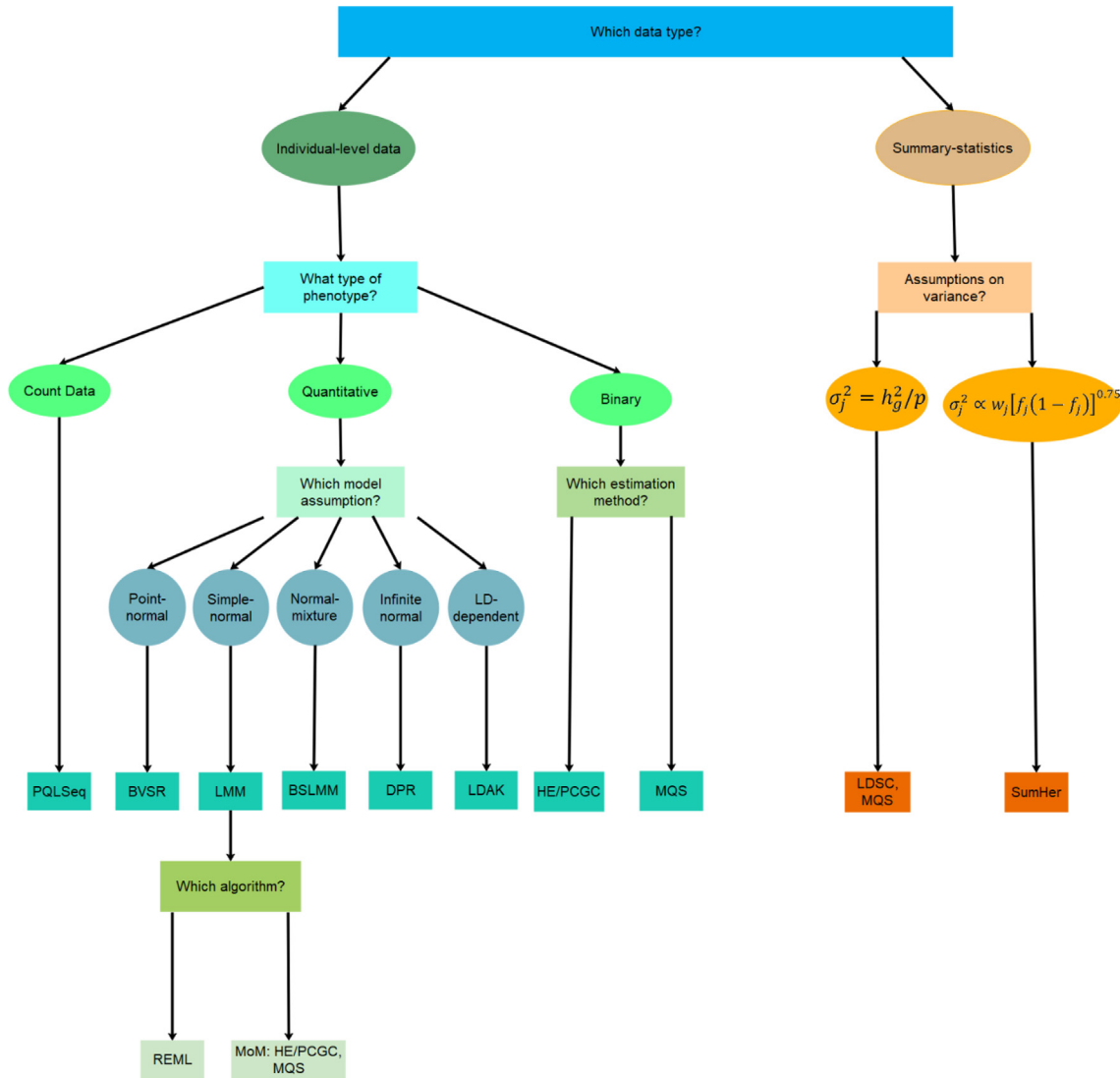


Fig. 3. A decision tree on what type of methods to use for SNP heritability estimation.

### 5. SNP heritability partitioning

In parallel to trait mapping efforts, large-scale functional genomic studies have yielded a rich source of SNP functional annotations [75–79]. Various discrete and continuous annotations are being developed to characterize the function of genetic variants [80–82]. For example, we can now classify genetic variants based on their genomic location (e.g., coding, intron and intergenic variants), role in protein structure and function (e.g., SIFT score [83] or PolyPhen score [84], ability to regulate gene expression (e.g., eQTL and ASE evidence [85–86], biochemical function (e.g., DNase I hypersensitive sites, metabolomic QTL evidence, and chromatin states [87–89], evolutionary significance (e.g., GERP score [90], and/or a combination of all these annotations (e.g., CADD score [76] and Eigen score [91]). These functional annotations are important predictors for SNP effects. Previous studies have shown that SNPs in certain functional categories (e.g., in promoters and enhancers) are more likely to be causal [92–93], tend to have larger effect sizes, and explain more heritability than SNPs in other categories (e.g., introns) [94–95]. Along with SNP properties (e.g., MAF and LD), incorporating SNP functional annotation is expected to improve SNP heritability estimation accuracy. In this section, we will introduce methods that are developed to estimate and parti-

tion SNP heritability by different SNP properties (e.g., GREML-MS and GREML-LDMS) or by different functional genomic annotations (e.g., stratified LDSC, MQS, and SMART).

Lee et al. [96] and Yang et al. [13] categorize SNPs into different categories based on MAFs and MAFs with LD scores, respectively, and assume the following extended LMM modeling assumption

$$\beta_j \sim N(0, \sigma_k^2), \tag{14}$$

if  $j$ -th SNP belongs to the  $k$ -th functional category. In this way, SNPs inside each functional category have their own variance component  $\sigma_k^2$ . When the SNPs are categorized based on their MAFs, GREML-MS implements the REML approach for fitting the extended LMM. When the SNPs are categorized based on both MAFs and LD scores, GREML-LDMS implements the REML approach for fitting the extended LMM. Incorporating MAFs and LD scores is particularly useful for analyzing whole-genome sequencing data that collect SNPs in high density with an excessive number of rare variants. Indeed, some of the methods mentioned in previous sections may yield inaccurate heritability estimates in different data types when their corresponding modeling assumptions do not fit the genetic architecture of the trait. For example, in whole genome sequencing data, a naïve application of LMM may lead to underestimation of

**Table 2**  
A summary of SNP heritability estimates for height using different methods.

References	Dataset	Data Type	Sample Size	Number of SNPs	SNP type (applicable AF)	Methods	SNP heritability Estimates
[12]	Australian data	Individual	35,189	294,831	Array (>0.01)	LMM/REML	0.449
[22]	Australian data	Individual	35,189	294,831	Array (>0.01)	BSLMM	0.41
						LMM/REML	0.42
						BVSR	0.15
[59]	Australian data	Individual	3,925	4,352,968	Imputed (>0.01)	MQS	0.28
						LMM/REML	0.27
						HE	0.25
						LDSC	0.21
[58]	Australian data	Individual	35,189	294,831	Array (>0.01)	PCGC/HE	0.537
						LMM/REML	0.510
[74]	24 Published GWAS	Summary	Average 121,000	4,555,718	Imputed (>0.01)	SumHer	0.46
						LDSC	0.20
[13]	UK10K	Individual	44,126	~17 M	Imputed (>0.0003)	GREML-LDMS	0.56
						GREML-MS	0.523

Table lists SNP heritability estimates for height reported in the previous literature. Columns contain the references where the SNP heritability estimates are reported (1st column), dataset name (2nd column), data type in terms of individual-level data versus summary statistics (3rd column), sample size (4th column), number of SNPs (5th column), genotype data type in terms of array data versus imputed data (6th column), used methods (7th column) and the SNP heritability estimates (8th column). Note that the heritability estimates for height in the Austrian data using the imputed data [59] is smaller than that using the array data, which seems to be general phenomenon for many other traits.

SNP heritability for traits whose underlying causal variants are mostly common. Instead, accurate SNP heritability estimation may require analysis in each MAF and/or LD stratum separately. To facilitate stratified analysis, GREML-MS stratifies the genetic variants based on their minor allele frequency into different MAF bins. That is, it estimates SNP heritability using SNPs in each MAF bin and then sums the estimates across bins. Similarly, GREML-LDMS stratifies the genetic variants based on both their MAF and LD and performs stratified heritability estimation. A previous study [13] had shown that, SNP heritability estimate for height was 52.3% by GREML-MS based on imputed data from 1,000 genome project reference. In the same data, SNP heritability estimate for height was 55.5% by using GREML-LDMS. Similar stratification ideas are applied for other methods, such as the stratified version of LD Score regression [97] or stratified LDSC, in which the SNP heritability is partitioned by functional annotations and estimated by using GWAS summary statistics. As introduced above, MQS [59] is based on a set of second moment matching equations determined by the MINQUE algorithm and has closed-form solutions for genetic variance components. MQS is also flexible in estimating heritability when SNPs are partitioned into different functional annotations. Hao et al. proposed SMART (Scalable Multiple Annotation integration for trait-Relevant Tissue) to mainly identify trait-relevant tissues by integrating multiple functional annotations jointly [98]. SMART modifies LMM to relate genetic effects with functional annotations by functionalizing the variant-specific variance components with respect to SNP annotations. The SNP heritability can be estimated based on the generalized estimation equation (GEE) that allows for only summary statistics.

## 6. Discussion

We have provided a technical review on a wide range of methods for SNP heritability estimation. We have focused on their modeling assumptions, their interconnected relationships, estimation algorithms, as well as their extensions towards different types of phenotypes and towards the use of summary statistics. Different methods have different benefits and may be preferred for heritability estimation of different traits or different data types. Indeed, heritability estimates for height in the literature vary depending on the particular methods used and depending on the datasets examined (Table 2). By detailing the technical properties of different methods, we hope that this review will serve as a useful reference for both methodologists who develop heritability

estimation methods and practitioners who perform heritability analyses.

As experimental technology develops and statistical methodology progresses, we are now able to achieve relatively robust and accurate SNP heritability estimation for many diseases and complex traits. For example, the estimates of SNP heritability for height is now above 50% [12,40]. However, these SNP heritability estimates are still less than that estimated from pedigree studies where the heritability of height is estimated to be 80%. This phenomenon,  $h_g^2 < h_{family}^2$  (SNP heritability < family-based heritability), is referred to as “still-missing heritability” [17]. Many explanations on “still-missing heritability” exist. Pedigree-based heritability estimation may be upward biased due to gene-environment interactions [5,6,14,15,19,99,100]. In contrast, inaccurate genotype calling in sequencing or array-based studies may lead to an underestimation of SNP heritability. Accurate heritability estimation requires the statistical modeling assumption to match the underlying genetic architecture, which depends on the minor allele frequency distribution of causal variants, LD pattern, and the strength of environmental components, all of which can be population specific. Consequently, SNP heritability estimates may change across populations and may change over time within each population [7]. A recent study reports that heterogeneity across sampling populations and time may contribute to part of the “still-missing heritability” [101] as most existing studies are carried out on individuals of European ancestry [12,59,72,74,98]. Using LMM models on datasets from seven sampling populations, this study discovered that at least 20% of missing heritability for BMI and 37% for years of education can be explained by individual heterogeneity. Therefore, understanding how various genetic, environmental, as well as study design factors influence the estimation of SNP heritability is an important future direction.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgement

This study was supported by the National Institutes of Health (NIH) Grants R01HG009124 and the National Science Foundation (NSF) Grant DMS1712933.

## References

- [1] Visscher PM, Hill WG, Wray NR. Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet* 2008;9(4):255–66.
- [2] Eaves LJ, Last KA, Young PA, Martin NG. Model-fitting approaches to the analysis of human behaviour. *Heredity* 1978;41(3):249.
- [3] Keller MC, Coventry WL. Quantifying and addressing parameter indeterminacy in the classical twin design. *Twin Research and Human Genetics*. 2005;8(3):201–13.
- [4] Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. *Nat Rev Genet* 2013;14(2):139.
- [5] J Mayhew A, Meyre D. Assessing the heritability of complex traits in humans: methodological challenges and opportunities. *Current genomics*. 2017;18(4):332–40.
- [6] Ober C, Vercelli D. Gene–environment interactions in human disease: nuisance or opportunity?. *Trends Genet* 2011;27(3):107–15.
- [7] Laird NM, Lange C. The fundamentals of modern statistical genetics. Springer Science & Business Media; 2010.
- [8] Rijdsdijk FV, Sham PC. Analytic approaches to twin data using structural equation models. *Briefings Bioinf* 2002;3(2):119–33.
- [9] Sham P, Cherny S, Purcell S. Application of genome-wide SNP data for uncovering pairwise relationships and quantitative trait loci. *Genetica* 2009;136(2):237–43.
- [10] Falconer DS. Introduction to quantitative genetics. Introduction to quantitative genetics. 1960.
- [11] Silventoinen K, Sarmalisto S, Perola M, Boomsma DI, Cornes BK, Davis C, et al. Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Res Human Genet* 2003;6(5):399–408.
- [12] Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010;42(7):565.
- [13] Yang J, Bakshi A, Zhu Z, Hemani G, Vinkhuyzen AA, Lee SH, et al. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet* 2015;47(10):1114.
- [14] Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, et al. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet* 2010;11(6):446.
- [15] Gibson G. Rare and common variants: twenty arguments. *Nat Rev Genet* 2012;13(2):135.
- [16] Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, Visscher PM. Pitfalls of predicting complex traits from SNPs. *Nat Rev Genet* 2013;14(7):507.
- [17] Witte JS, Visscher PM, Wray NR. The contribution of genetic variants to disease depends on the ruler. *Nat Rev Genet* 2014;15(11):765.
- [18] Yang J, Zeng J, Goddard ME, Wray NR, Visscher PM. Concepts, estimation and interpretation of SNP-based heritability. *Nat Genet* 2017;49(9):1304.
- [19] Timpson NJ, Greenwood CM, Soranzo N, Lawson DJ, Richards JB. Genetic architecture: the shape of the genetic contribution to human traits and disease. *Nat Rev Genet* 2018;19(2):110.
- [20] Evans LM, Tahmasbi R, Vrieze SI, Abecasis GR, Das S, Gazal S, et al. Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nat Genet* 2018;50(5):737.
- [21] Strandén I, Christensen OF. Allele coding in genomic evaluation. *Genet Select Evol* 2011;43(1):25.
- [22] Zhou X, Carbonetto P, Stephens M. Polygenic modeling with Bayesian sparse linear mixed models. *PLoS Genet* 2013;9(2):e1003264.
- [23] de los Campos G, Vazquez AI, Fernando R, Klimentidis YC, Sorensen D. Prediction of complex human traits using the genomic best linear unbiased predictor. *PLoS genetics*. 2013;9(7).
- [24] Visscher PM. Sizing up human height variation. *Nat Genet* 2008;40(5):489.
- [25] Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, et al. Genome-wide association analysis identifies 20 loci that influence adult height. *Nat Genet* 2008;40(5):575.
- [26] Lettre G, Jackson AU, Gieger C, Schumacher FR, Berndt SI, Sanna S, et al. Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat Genet* 2008;40(5):584.
- [27] Gudbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, et al. Many sequence variants affecting diversity of adult human height. *Nat Genet* 2008;40(5):609.
- [28] Yi N, Xu S. Bayesian LASSO for quantitative trait loci mapping. *Genetics* 2008;179(2):1045–55.
- [29] Hoggart CJ, Whittaker JC, De Iorio M, Balding DJ. Simultaneous analysis of all SNPs in genome-wide and re-sequencing association studies. *PLoS Genet* 2008;4(7):e1000130.
- [30] Wu TT, Chen YF, Hastie T, Sobel E, Lange K. Genome-wide association analysis by lasso penalized logistic regression. *Bioinformatics* 2009;25(6):714–21.
- [31] Logsdon BA, Hoffman GE, Mezey JG. A variational Bayes algorithm for fast and accurate multiple locus genome-wide association analysis. *BMC Bioinf* 2010;11(1):58.
- [32] Guan Y, Stephens M. Bayesian variable selection regression for genome-wide association studies and other large-scale problems. *Annals Appl Statist* 2011;5(3):1780–815.
- [33] Carbonetto P, Stephens M. Scalable variational inference for Bayesian variable selection in regression, and its accuracy in genetic association studies. *Bayesian Anal* 2012;7(1):73–108.
- [34] Pankow JS, Folsom AR, Cushman M, Borecki IB, Hopkins PN, Eckfeldt JH, et al. Familial and genetic determinants of systemic markers of inflammation: the NHLBI family heart study. *Atherosclerosis*. 2001;154(3):681–9.
- [35] Lange LA, Burdon K, Langefeld CD, Liu Y, Beck SR, Rich SS, et al. Heritability and expression of C-reactive protein in type 2 diabetes in the Diabetes Heart Study. *Ann Hum Genet* 2006;70(6):717–25.
- [36] Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Human Genet*. 2011;88(1):76–82.
- [37] Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* 2012;44(7):821.
- [38] Wu Z, Sun Y, He S, Cho J, Zhao H, Jin J. Detection boundary and higher criticism approach for rare and weak genetic effects. *Ann Appl Statist* 2014;8(2):824–51.
- [39] Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell* 2017;169(7):1177–86.
- [40] Zhu X, Stephens M. Bayesian large-scale multiple regression with summary statistics from genome-wide association studies. *Ann Appl Statist* 2017;11(3):1561.
- [41] Speed D, Hemani G, Johnson MR, Balding DJ. Improved heritability estimation from genome-wide SNPs. *Am J Human Genet* 2012;91(6):1011–21.
- [42] Speed D, Cai N, Johnson MR, Nejentsev S, Balding DJ, Consortium U. Reevaluation of SNP heritability in complex human traits. *Nature genetics*. 2017;49(7):986.
- [43] Hayes B, Goddard M. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 2001;157(4):1819–29.
- [44] Verbyla KL, Bowman PJ, Hayes BJ, Goddard ME, editors. Sensitivity of genomic selection to using different prior distributions. *BMC proceedings*; 2010: BioMed Central.
- [45] Verbyla KL, Hayes BJ, Bowman PJ, Goddard ME. Accuracy of genomic selection using stochastic search variable selection in Australian Holstein Friesian dairy cattle. *Genet Res* 2009;91(5):307–11.
- [46] Hayes BJ, Pryce J, Chamberlain AJ, Bowman PJ, Goddard ME. Genetic architecture of complex traits and accuracy of genomic prediction: coat colour, milk-fat percentage, and type in Holstein cattle as contrasting model traits. *PLoS Genet* 2010;6(9):e1001139.
- [47] Habier D, Fernando RL, Kizilkaya K, Garrick DJ. Extension of the Bayesian alphabet for genomic selection. *BMC Bioinf* 2011;12(1):186.
- [48] Park T, Casella G. The Bayesian lasso. *J Am Stat Assoc* 2008;103(482):681–6.
- [49] Moser G, Lee SH, Hayes BJ, Goddard ME, Wray NR, Visscher PM. Simultaneous discovery, estimation and prediction analysis of complex traits using a Bayesian mixture model. *PLoS Genet* 2015;11(4):e1004969.
- [50] Zeng J, De Vlaming R, Wu Y, Robinson MR, Lloyd-Jones LR, Yengo L, et al. Signatures of negative selection in the genetic architecture of human complex traits. *Nat Genet* 2018;50(5):746.
- [51] Zeng P, Zhou X. Non-parametric genetic prediction of complex traits with latent Dirichlet process regression models. *Nat Commun* 2017;8(1):456.
- [52] Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet* 2015;47(9):1091.
- [53] Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet* 2016;48(3):245.
- [54] Falconer DS. The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet* 1965;29(1):51–76.
- [55] Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *Am J Human Genet* 2011;88(3):294–305.
- [56] Dempster ER, Lerner IM. Heritability of threshold characters. *Genetics* 1950;35(2):212.
- [57] Chen G-B. Estimating heritability of complex traits from genome-wide association studies using IBS-based Haseman-Elston regression. *Front Genet* 2014;5:107.
- [58] Golan D, Lander ES, Rosset S. Measuring missing heritability: inferring the contribution of common variants. *Proc Natl Acad Sci* 2014;111(49):E5272–81.
- [59] Zhou X. A unified framework for variance component estimation with summary statistics in genome-wide association studies. *Ann Appl Statist* 2017;11(4):2027.
- [60] Weissbrod O, Flint J, Rosset S. Estimating SNP-based heritability and genetic correlation in case-control studies directly and with summary statistics. *Am J Human Genet* 2018;103(1):89–99.
- [61] Tung J, Zhou X, Alberts SC, Stephens M, Gilad Y. The genetic architecture of gene expression levels in wild baboons. *Elife* 2015;4:e04729.
- [62] Price AL, Helgason A, Thorleifsson G, McCarroll SA, Kong A, Stefansson K. Single-Tissue and Cross-Tissue Heritability of Gene Expression Via Identity-by-Descent in Related or Unrelated Individuals. *Plos Genet*. 2011;7(2).
- [63] Wright FA, Sullivan PF, Brooks AI, Zou F, Sun W, Xia K, et al. Heritability and genomics of gene expression in peripheral blood. *Nat Genet*. 2014;46(5):430–7.
- [64] Monks SA, Leonardson A, Zhu H, Cundiff P, Pietrusiak P, Edwards S, et al. Genetic inheritance of gene expression in human cell lines. *Am J Hum Genet*. 2004;75(6):1094–105.

- [65] Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, et al. Genetics of gene expression and its effect on disease. *Nature* 2008;452(7186):423–U2.
- [66] Banovich NE, Lan X, McVicker G, van de Geijn B, Degner JF, Blischak JD, et al. Methylation QTLs Are Associated with Coordinated Changes in Transcription Factor Binding, Histone Modifications, and Gene Expression Levels. *PLoS Genet*. 2014;10(9):e1004663.
- [67] McRae AF, Powell JE, Henders AK, Bowdler L, Hemani G, Shah S, et al. Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biol*. 2014;15(5):R73.
- [68] Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, et al. Epigenome-Wide Scans Identify Differentially Methylated Regions for Age and Age-Related Phenotypes in a Healthy Ageing Population. *PLoS Genet*. 2012;8(4):189–200.
- [69] Cheng CS, Gate RE, Aiden AP. Genetic determinants of co-accessible chromatin regions in T cell activation across humans. *BioRxiv*. 2017.
- [70] Wheeler HE, Shah KP, Brenner J, Garcia T, Aquino-Michaels K, Cox NJ, et al. Survey of the Heritability and Sparse Architecture of Gene Expression Traits across Human Tissues. *PLoS Genet*. 2016;12(11):e1006423.
- [71] Sun S, Zhu J, Mozaffari S, Ober C, Chen M, Zhou X. Heritability estimation and differential analysis of count data with generalized linear mixed models in genomic sequencing studies. *Bioinformatics* 2018;35(3):487–96.
- [72] Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Patterson N, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015;47(3):291.
- [73] Robbins H, Monro S. A stochastic approximation method. *The annals of mathematical statistics*. 1951:400–7.
- [74] Speed D, Balding DJ. SumHer better estimates the SNP heritability of complex traits from summary statistics. *Nat Genet* 2019;51(2):277.
- [75] Genomes Project C, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491(7422):56–65.
- [76] Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46(3):310–5.
- [77] Consortium GT. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015;348(6235):648–60.
- [78] Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. *Nature* 2015;518(7539):317–30.
- [79] Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;489(7414):57.
- [80] Carithers LJ, Moore HM. The Genotype-Tissue Expression (GTEx) Project. *Biopreserv Biobanking* 2015;13(5):307–8.
- [81] Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature* 2015;518(7539):331–6.
- [82] Kellis M, Wold B, Snyder MP, Bernstein BE, Kundaje A, Marinov GK, et al. Defining functional DNA elements in the human genome. *PNAS* 2014;111(17):6131–8.
- [83] Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4(7):1073–81.
- [84] Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics / editorial board, Jonathan L Haines [et al]*. 2013;Chapter 7:Unit7 20.
- [85] Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature* 2010;464(7289):768–72.
- [86] Tung J, Zhou X, Alberts SC, Stephens M, Gilad Y. The genetic architecture of gene expression levels in wild baboons. *eLife*. 2015;4.
- [87] Pique-Regi R, Degner JF, Pai AA, Gaffney DJ, Gilad Y, Pritchard JK. Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. *Genome Res* 2011;21(3):447–55.
- [88] Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 2012;9(3):215–6.
- [89] McVicker G, van de Geijn B, Degner JF, Cain CE, Banovich NE, Raj A, et al. Identification of Genetic Variants That Affect Histone Modifications in Human Cells. *Science* 2013;342(6159):747–9.
- [90] Cooper GM, Stone EA, Asimenos G, Program NCS, Green ED, Batzoglu S, et al. Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res* 2005;15(7):901–13.
- [91] Ionita-Laza I, McCallum K, Xu B, Buxbaum JD. A spectral approach integrating functional genomic annotations for coding and noncoding variants. *Nat Genet* 2016;48(2):214–20.
- [92] Pickrell JK. Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am J Hum Genet* 2014;94(4):559–73.
- [93] Schork AJ, Thompson WK, Pham P, Torkamani A, Roddey JC, Sullivan PF, et al. All SNPs are not created equal: genome-wide association studies reveal a consistent pattern of enrichment among functionally annotated SNPs. *PLoS Genet* 2013;9(4):e1003449.
- [94] Gusev A, Lee SH, Trynka G, Finucane H, Vilhjalmsson BJ, Xu H, et al. Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. *Am J Hum Genet* 2014;95(5):535–52.
- [95] Kichaev G, Yang WY, Lindstrom S, Hormozdiari F, Eskin E, Price AL, et al. Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLoS Genet* 2014;10(10):e1004722.
- [96] Lee SH, Yang J, Chen G-B, Ripke S, Stahl EA, Hultman CM, et al. Estimation of SNP heritability from dense genotype data. *Am J Human Genet* 2013;93(6):1151–5.
- [97] Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh P-R, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* 2015;47(11):1228.
- [98] Hao X, Zeng P, Zhang S, Zhou X. Identifying and exploiting trait-relevant tissues with multiple functional annotations in genome-wide association studies. *PLoS Genet* 2018;14(1):e1007186.
- [99] van Rheenen W, Peyrot WJ, Schork AJ, Lee SH, Wray NR. Genetic correlations of polygenic disease traits: from theory to practice. *Nat Rev Genet* 2019;1.
- [100] Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461(7265):747.
- [101] Tropf FC, Lee SH, Verweij RM, Stulp G, Van Der Most PJ, De Vlaming R, et al. Hidden heritability due to heterogeneity across seven populations. *Nat Hum Behav* 2017;1(10):757–65.