Partitioning Heritability of Regulatory and Cell-Type-Specific Variants across 11 Common Diseases

Alexander Gusev,[1,*] S. Hong Lee,[2] Gosia Trynka,[3,4,5,6,16] Hilary Finucane,[7] Bjarni J. Vilhjálmsson,[1] Han Xu,[8] Chongzhi Zang,[8] Stephan Ripke,[9,10] Brendan Bulik-Sullivan,[9,10] Eli Stahl,[11] Schizophrenia Working Group of the Psychiatric Genomics Consortium, SWE-SCZ Consortium, Anna K. Kähler,[12] Christina M. Hultman,[12] Shaun M. Purcell,[9,10,11] Steven A. McCarroll,[10] Mark Daly,[6,9,10] Bogdan Pasaniuc,[13] Patrick F. Sullivan,[12,14] Benjamin M. Neale,[6,9,10] Naomi R. Wray,[2] Soumya Raychaudhuri,[3,4,5,6,15] and Alkes L. Price[1,6,*]

Regulatory and coding variants are known to be enriched with associations identified by genome-wide association studies (GWASs) of complex disease, but their contributions to trait heritability are currently unknown. We applied variance-component methods to imputed genotype data for 11 common diseases to partition the heritability explained by genotyped SNPs (h_g^2) across functional categories (while accounting for shared variance due to linkage disequilibrium). Extensive simulations showed that in contrast to current estimates from GWAS summary statistics, the variance-component approach partitions heritability accurately under a wide range of complex-disease architectures. Across the 11 diseases DNaseI hypersensitivity sites (DHSs) from 217 cell types spanned 16% of imputed SNPs (and 24% of genotyped SNPs) but explained an average of 79% (SE = 8%) of h_g^2 from imputed SNPs (5.1× enrichment; p = 3.7 × 10^{-17}) and 38% (SE = 4%) of h_g^2 from genotyped SNPs (1.6× enrichment, p = 1.0 × 10^{-4}). Further enrichment was observed at enhancer DHSs and cell-type-specific DHSs. In contrast, coding variants, which span 1% of the genome, explained <10% of h_g^2 despite having the highest enrichment. We replicated these findings but found no significant contribution from rare coding variants in independent schizophrenia cohorts genotyped on GWAS and exome chips. Our results highlight the value of analyzing components of heritability to unravel the functional architecture of common disease.

Introduction

Recent work by ENCODE and other projects[1,2] has shown that specific classes of variants can have complex and diverse impacts on cell function and phenotype.[3–10] Although the importance of coding variation has long been understood, these projects identified other genomic regions that can contribute to function and highlighted the role of regulatory variants. With many potentially informative functional categories and competing biological hypotheses, quantifying the contribution of variants in these categories to heritability of complex traits would inform trait biology and focus genetic mapping.

The availability of significantly associated variants from hundreds of genome-wide association studies (GWASs)[11] has opened one avenue for quantifying enrichment. Indeed, 11% of GWAS hits lie in coding regions,[11] 57% of noncoding GWAS hits lie in broadly defined DNaseI hypersensitivity sites (DHSs; spanning 42% of the genome),[3] and still additional GWAS hits tag these regions. The full distribution of GWAS association statistics exhibits enriched p values in coding regions and UTRs.[12] Analysis of DHS subclasses and other histone marks has revealed a complex pattern of cell-type-specific relationships with known disease associations.[6] Recent work has also shown that functional enrichment can be leveraged for increasing association mapping power.[13]

Although relative enrichment has been documented, the question of how much each category contributes to disease heritability remains unanswered.[14,15] Recently, investigators have used variance-component methods to estimate the total additive variance explained by all genotyped SNPs (h_g^2),[16,17] and to estimate the h_g^2 of many quantitative and dichotomous traits.[18–22] We propose joint estimation of h_g^2 from functional-category-specific variance components for assessing enrichment. In contrast to analyses of top GWAS hits, the variance-component approach leverages the entire polygenic architecture of each trait and accounts for pervasive linkage disequilibrium (LD) across functional categories. Indeed, our

simulations showed that this approach provides accurate genome-wide estimates of functional enrichment in diverse genetic architectures. We applied variance-component methods to functional categories in GWAS- and exome-chip data from over 100,000 samples in 11 traits.

Material and Methods

Estimating Enrichment of h_g^2 with Variance Components

For a single component of genotyped (or imputed) SNPs, we define h_g^2, an underlying parameter in the population, as the r^2 between the true phenotype and the best linear prediction over those SNPs. With multiple components, the goal of the partitioned analysis is to quantify the h_g^2 directly explained by SNPs in each functional category while excluding tagging of SNPs in other categories. We thus define the h_g^2 for each functional category as the r^2 between the true phenotype and the prediction only from SNPs in that functional category when all functional categories are jointly analyzed for a best linear prediction. When SNPs are in LD, this definition remains valid as long as the individual causal effect sizes are independent, as we would expect in highly polygenic traits. For disease traits, we model the phenotype (and corresponding h_g^2) by using the liability-threshold model, in which individuals whose underlying unobserved continuous liability exceeds a threshold are labeled as disease case subjects.[19,23]

We estimate h_g^2 jointly across multiple variance components, each constructed from variants belonging to nonoverlapping functional categories. The underlying model assumes that SNP effect sizes are drawn from a normal distribution with category-specific variance. (We note that the normality assumption is unrealistic; previous work in the single-variance-component case has indicated that this does not introduce bias, although modeling a more realistic mixture distribution can increase precision.[24] Because of computational constraints, we do not consider mixture distributions here.) The model relates the observed phenotypic covariance to a weighted sum of genetic relationship matrices computed from SNPs in each category. The joint estimate allows all components to compete for shared variance due to LD.

Formally, for a functional categories each containing the set of SNPs S_i (of size M_i), we model the phenotype as a sum of individual SNP effect sizes:

$$y = \sum_{i=1}^{a} \sum_{s \in S_i} W_s \beta_s^i + e,$$

where W_s is the genotype at SNP s, β_s^i is the effect size at SNP s in category i and is drawn from category-specific normal distribution $\beta^i \sim N(0, \sigma_i^2)$, and e is the residual effect $e \sim N(0, \sigma_e^2)$. We assume that for each annotation i, SNPs normalized to have mean 0 and variance 1 are contained in the matrix W_i. The variance of the phenotype is then modeled as

$$V(y) = \sum_{i=1}^{a} K_i \sigma_i^2 + e,$$

where each K_i represents a genetic-relationship matrix (GRM) computed directly from the SNPs in annotation i as

$$K_i = W_i W_i' / M_i.$$

The corresponding σ are then jointly inferred with the REML algorithm in GCTA (Genome-wide Complex Trait Analysis),[16,17] yielding

$$h_{gi}^2 = \frac{\sigma_{gi}^2}{\sum_{j=1}^{a} \sigma_{gi}^2 + \sigma_e^2}.$$

The inverse of the final average-information matrix yields an estimate of the corresponding error-covariance matrix of the variance-component estimates.[25] We use the error-covariance matrix and delta method[26] to compute SEs on h_g^2 and the percentage of h_g^2 while accounting for error correlations (referred to here as the analytical SE[27]). All estimates of h_g^2 were transformed to the liability scale[19] with the prevalence values in Table S1 (available online). We evaluated the accuracy of the analytical SE for both quantitative and ascertained traits and found it to correspond well to the true SD under reasonable polygenicity (see Appendix A). Meta-analysis estimates were computed with inverse-variance weighting:[28] given individual study estimates h_{gi}^2, analytical SE_i, and corresponding weight $w_i = 1/SE_i^2$, the meta-analysis mean is equal to

$$\frac{\sum_i w_i \times h_{gi}^2}{\sum_i w_i},$$

and the meta-analysis SE is equal to $\sqrt{1/\sum_i w_i}$.

Enrichment is computed for each category i as the ratio of the percentage of h_g^2 (the percentage of h_{gi}^2 in category i) to the percentage of SNP_i (the percentage of SNPs in category i) and is tested for significance by Z score relative to a null of 1:0 with the (likewise-rescaled) analytical SE. Under the assumption that all causal variants are typed, this statistic is equivalent to the relative risk that a SNP in category i is causal in comparison to an average SNP. To achieve unbiasedness, the estimate of h_g^2 is not constrained to lie inside the plausible 0–1 bound, which can lead to negative estimates in rare instances.

Estimating Enrichment from Summary Statistics

We considered alternative methods for estimating functional enrichment from summary association statistics. The simplest approach is to directly count the number of individual genome-wide-significant variants in each functional annotation and compare to the null expectation from all SNPs (or random SNPs matched on certain features). This approach can either include all significant markers or restrict to the most significant variant in each locus. The genome-wide-significant-SNP approach has been extended to the full distribution of association statistics for quantifying overall p value enrichment.[3] Over increasingly restrictive p value thresholds, the fraction of SNPs passing a given threshold and belonging to each category is computed and normalized by the category-specific genome-wide fraction. The distributions are then inspected visually for enrichment or assessed by permutation. For completeness, we considered two additional methods—stratified quantile-quantile (Q-Q) plots[12] and Bayesian hierarchical modeling (fgwas)[13]—which assess functional enrichment but are primarily focused on improving association mapping power (see Discussion).

Data Sets Analyzed

11 Diseases from WTCCC1 and WTCCC2
We analyzed seven traits from Wellcome Trust Case Control Consortium 1 (WTCCC1) and four traits from WTCCC2 for a total

47,000 samples (Table S1). Estimates of h_g^2 are particularly sensitive to individually small artifacts or batch effcts,[19,29] and we followed the rigorous quality-control (QC) protocol outlined previously[21] by removing any SNPs that were below a minor allele frequency (MAF) of 0.01, were above 0.002 missingness, or deviated from Hardy-Weinberg equilibrium at a p value below 0.01. For each case-control cohort, we removed SNPs that had differential missingness with a p value below 0.05. We excluded one of any pair of samples with kinship entries ≥ 0.05[19] and performed five rounds of outlier removal whereby all individuals more than 6 SDs away from the mean along any of the top 20 eigenvectors were removed and all eigenvectors were recomputed[30] (Figure S1). For all autoimmune diseases analyzed (rheumatoid arthritis [RA], Crohn disease [CD], type 1 diabetes [T1D], ulcerative colitis [UC], multiple sclerosis [MS], and ankylosing spondylitis [AS]), we also excluded from the analysis any SNPs in the well-studied major histocompatibility complex (MHC) locus (chr6: 26–34 Mb), which is known to have a complex LD structure, and many heterogeneous variants of strong effect for these traits.

The WTCCC1 samples were phased and imputed as described in Gusev et al.[21] The WTCCC2 samples were split into two cohorts by platform, and each cohort was imputed separately according to the following protocol. All samples in a cohort were phased together in 10 Mb blocks with HAPI-UR (Haplotype Inference for Unrelated Samples)[31] (see Web Resources) and three rounds of phasing and consensus voting. All phased samples in a cohort were then imputed in 1 Mb blocks with IMPUTE2[32] (see Web Resources) and the 1000 Genomes[33] Phase I integrated haplotypes (September 2013 release; see Web Resources) with no singletons. Where relevant, the Oxford recombination map[34] was used. Markers with an information (info) score greater than 0.5 were retained. Finally, SNPs were excluded if they met any of the following criteria in any case or control population: Hardy-Weinberg p value < 0.05, per-locus missingness > 0.05, MAF < 0.01, or case-control differential missingness p value < 0.05.

Schizophrenia Cohort from the Psychiatric Genomics Consortium
We analyzed 24,926 schizophrenia (SP) subjects and 33,271 control individuals from 33 cohorts from the Psychiatric Genomics Consortium (PGC2); they were typed on a variety of platforms, quality controlled, and imputed to the 1000 Genomes reference panel as previously described[35] (Tables S1 and S2). Because of computational constraints, we split the cohort into four subgroups of individuals typed on similar platforms; each contained roughly 10,000–20,000 samples. We performed all analyses on the intersection of well-imputed SNPs within each subgroup, ranging from four to five million, and reported meta-analyzed results. Individual study identifiers and 20 multidimensional-scaling components were included as fixed-effect covariates in all analyses.

Swedish SP Exome Chip
We analyzed 12,674 Swedish samples typed on GWAS and exome chips (Tables S1 and S3). The exome chip yielded 238,652 SNPs (including monomorphic sites), of which 10,567 were also typed on a mix of Affymetrix GWAS chips (exome-chip calls were retained). The GWAS-chip data contained an intersection of 163,051 SNPs typed on all platforms in addition to per-platform imputation from 1000 Genomes for a total of 5,053,934 SNPs imputed on all platforms. Principal-component analysis (PCA) of the GWAS data revealed a large cluster of "homogenous" Swedish samples and clines related to Northern Swedish and Finnish admixture (Figure S2). After excluding samples that (1) were not typed on both GWAS and exome chips, (2) failed QC, (3) were PCA outliers by 6 SDs, or (4) were in a pair with GRM values >

0.05 (close relatives), we retained a total of 8,967 samples, of which 6,375 were of "homogenous" Swedish ancestry. In all of our analyses, rare variants had a MAF < 0.01, and common variants had a MAF $\geq$ 0.01. Simulations were performed on the homogenous samples (without principal components). We performed analyses of real phenotypes on the homogeneous samples and included the top 20 principal components as covariates (to account for any residual population structure; analyses on the full cohort are reported in Tables S23 and S25).

Functional Annotations
We annotated the genome by using six primary categories (Table S4): (1) coding, (2) UTR, (3) promoter, (4) DHS in any of 217 cell types, (5) intronic, and (6) intergenic. Each SNP was then assigned a unique annotation defined by the first of these categories with which it was annotated, resulting in six nonoverlapping variance components (the DHS category was thus restricted to distal regions). Each resulting category exhibited similar average allele frequency and imputation accuracy, although the DHS category had systematically lower LD[36] (Table S5). We also computed the "effective" number of SNPs in each category by using an LD-based metric that does not depend on sample size.[36,37] Table S6 shows that this metric was not substantially different from the actual percentage of SNPs used in imputed data, given that DHSs harbored slightly more effective SNPs (15.7% SNPs versus 18.9% effective SNPs) as a result of lower LD. For the imputed categories analyzed here, the differences in the percentage of SNPs, percentage of effective SNPs, and percentage of physical size were relatively minor. A greater difference was observed for genotyped SNPs: 23.6% of DHS SNPs corresponded to 33.6% of effective SNPs, suggesting that DHS enrichments from genotyped data might be indicative of better tagging.

For the DHS annotation, we used DNase sequencing libraries downloaded from ENCODE and Epigenome Roadmap projects in May 2012 and merged biological replicates into a single library (GEO accession numbers are available in Table S7). We used BOWTIE v.1.0[38] to align raw read sequences to UCSC Genome Browser hg19 and used MACS v.2.0 with false-discovery rate < 0.01 (the default cutoff) and Benjamini–Hochberg correction[39] to call DHS peaks. For the primary analysis, all peaks were merged into a single DHS annotation spanning 16% of the genome. We note that 98% of the primary DHS annotation was covered by the DHSs released by Maurano et al.[3] (spanning 37% of the genome), and 67% of the primary DHS annotation was covered by the DHSs analyzed in Thurman et al.[4] (spanning 15% of the genome). For the cell-type-specific analysis, duplicate lines were merged to form a final set of 83 unique cell types. The resulting annotations are available for download (see Web Resources).

Segway-chromHMM combined genome segmentations[40] were downloaded for six cell lines (see Web Resources). All regions classified as enhancers or weak enhancers were then combined into a single enhancer annotation. DNaseI digital genomic footprinting (DGF) regions were downloaded for 57 cell lines (see Web Resources). All regions from the narrow-peak classification were then merged into a single DGF annotation.

Simulation Framework
The goal of our simulations was to demonstrate that the partitioned h_g^2 properly recovers the heritability explained by causal variants in a given functional category under a variety of disease architectures. We performed simulations in genotyped and

Variance components
No enrichment

p value enrichment
No enrichment
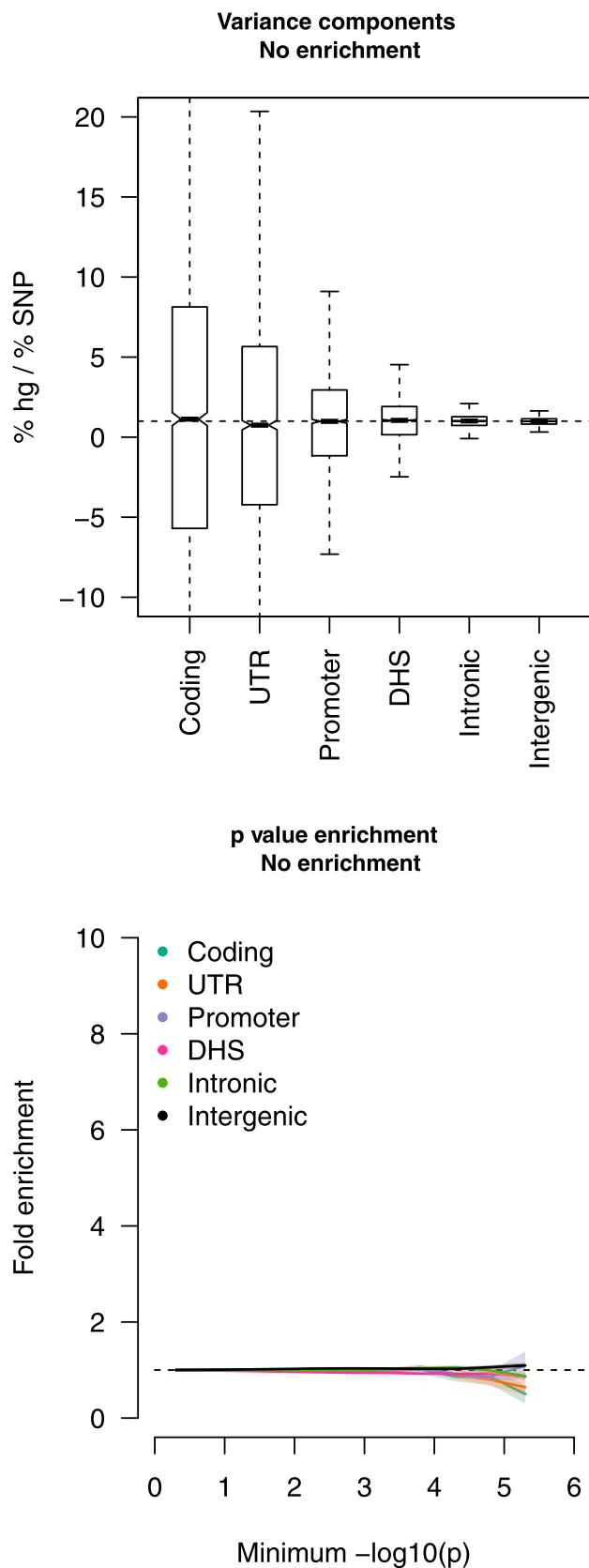
Figure 1. Estimates of Functional Enrichment under the Null
We simulated a polygenic disease architecture in imputed data with no functional enrichment (see text). Simulated phenotypes were tested with the variance-component method (top) from 3,000 simulations and with p value enrichment (bottom) from 100 simulations. In the variance-component subplot, the thin line represents the median, boxes represent the first and third quartiles, and whiskers represent the 1.5× interquartile range from the first to the third quartile. A subplot of p value enrichment shows 1.96× SE as shaded regions.

imputed data in 4,414 samples from the WTCCC1 coronary artery disease (CAD) case-control cohort together with the six main functional annotations to evaluate robustness and accuracy of the proposed variance-component method and the p-value-enrichment approach; we note that the genome-wide-significant-SNP approach is subsumed by the latter and is not reported separately in most analyses. For each simulation, 10% of the (genotyped or imputed) SNPs were randomly sampled to be causal, and normally distributed effect sizes were assigned to each SNP such that each explained equal variance in expectation. Additive phenotypes were then constructed, and random noise was added for an overall h_g^2 of 0.50. Except when evaluating h_g^2 between genotyped and imputed SNPs, we did not hide causal variants from the analyses, corresponding to the assumption that all causal variants are typed. We evaluated the variance-component model by using multiple components with GCTA in the unconstrained mode. For approaches based on summary statistics, we computed Z scores, SEs, and p values for the univariate regression of each SNP to a simulated phenotype.

Results

Simulations

We first evaluated the calibration of the methods in simulations of no enrichment by assuming a MAF-independent architecture where causal variants were uniformly sampled from the genome (see Material and Methods). We observed no significant deviations from the null for any categories estimated by variance components or p value enrichment (Figure 1). To evaluate possible biases due to MAF-dependent architectures,[21,41,42] we also considered a low-frequency architecture where only SNPs with a MAF below 0.05 can be causal and a DHS-low-frequency architecture where causal DHS variants are drawn from MAF below 0.05 and all other variants are drawn from any MAF (Figures S3 and S4). Results were generally similar to the MAF-independent architecture, although variance-component estimates exhibited slight but statistically significant deviations for the promoter and UTR categories, which were very small and in tight LD with each other.

We next considered simulations with maximal enrichment, where all causal variants were drawn from a single functional category. MAF-independent results for the coding and DHS categories are shown in Figure 2 (see Figure S5 for other results). The variance-component estimate of the percentage of h_g^2 was again around 100% for the true causal category and 0% for all others. The plots of p value enrichment correctly demonstrated significant enrichment for five of the categories, but not the DHS category, which, when causal, was not significantly different from the null. This lack of enrichment at DHSs and not at other large categories was most likely due to the uniquely lower
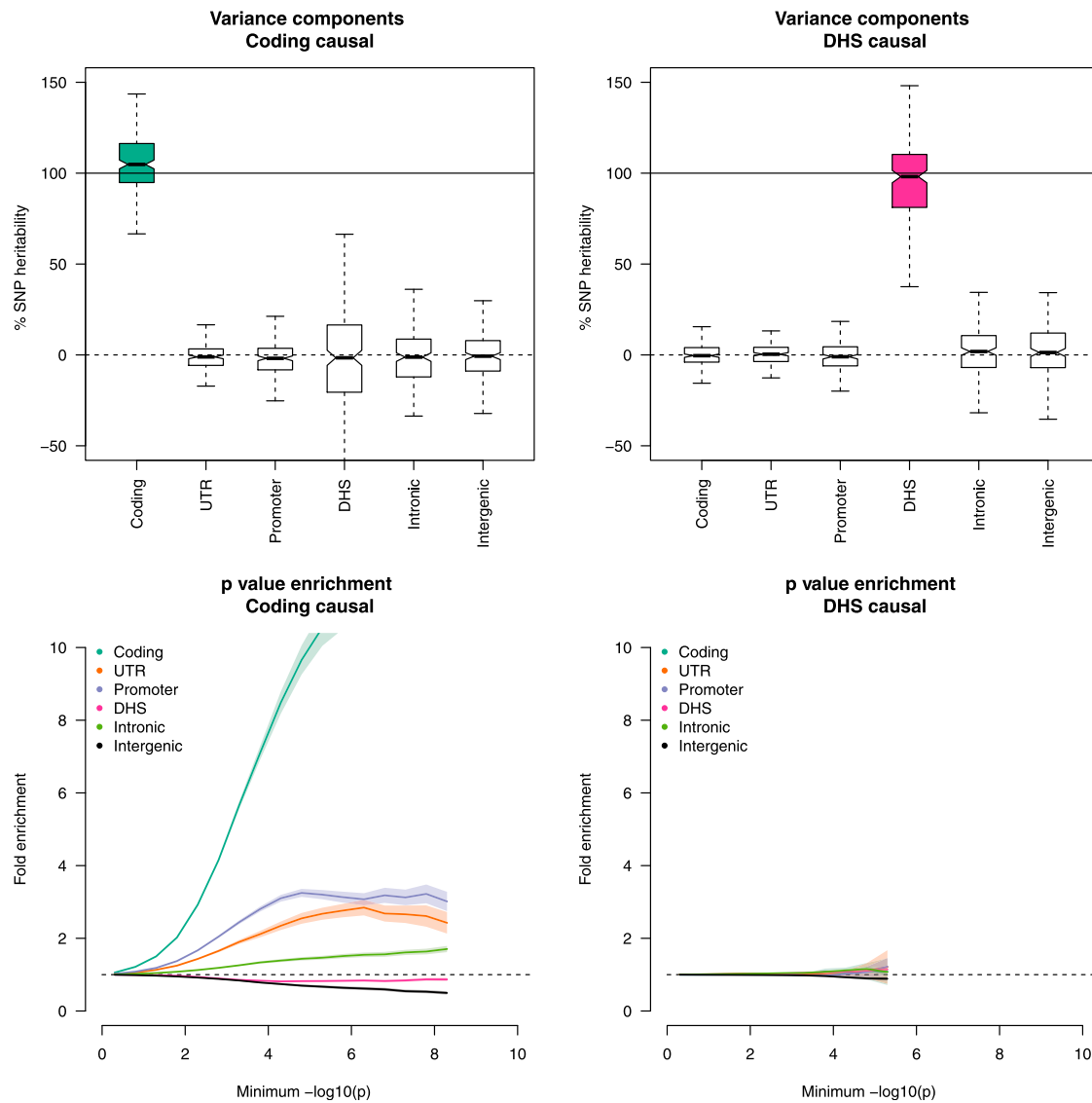
Figure 2. Estimates of Functional Enrichment from a Single Causal Category
We simulated a polygenic disease architecture in imputed data with causal SNPs drawn from a single functional category, corresponding to complete enrichment. Simulated phenotypes were tested with the variance-component method (top) from 200 simulations and with p value enrichment (bottom) from 100 simulations. In the variance-component subplot, the thin line represents the median, boxes represent the first and third quartiles, and whiskers represent the 1.5× interquartile range from the first to the third quartile. Subplots of p value enrichment show 1.96× SE across simulations as shaded regions. For each method, only the coding-causal and DHS-causal scenarios are shown (additional simulations appear in Figures S6 and S7).

LD of DHS SNPs (Table S5). For the small categories (coding, UTR, and promoter), true causals in one category always yielded false p value enrichment in the others because of their close proximity and high LD (Figure 2; Figures S6 and S7). In the MAF-dependent scenarios, the variance-component estimate of h_g^2 was nearly unbiased: it had slight but significant inflation at the coding and UTR categories when they contained 100% of h_g^2 (Figure S8). Plots of p value enrichment exhibited similar patterns as in the MAF-independent simulations, and the DHS category was further falsely depleted (Figure S7).

To investigate the differences between genotype- and imputation-based estimates, we partitioned h_g^2 of category-specific phenotypes simulated from imputed SNPs by using components constructed from genotyped SNPs only. If the genotypes are reasonable proxies for imputed variants, 100% of h_g^2 should again be partitioned into each truly causal category. Instead, we observed significant deviations for all of the categories, and h_g^2 was partitioned into nearby categories as a result of incomplete tagging (Figures S9 and S10). In particular, less than half of the h_g^2 at imputed DHSs was partitioned into the DHS category in genotype data. Thus, estimates produced with only genotyped SNPs can severely underestimate enrichment. The difference between genotyped and imputed simulations suggests that estimates from imputed SNPs could also underestimate the true enrichments or depletions for rare causal variants that are absent from 1000 Genomes or are

poorly imputed. We investigated this possibility by using the exome-chip SP data (see below). We separately assessed the impact of imputation error by simulating phenotypes with induced genotype noise proportional to the per-SNP imputation quality score (info score; Supplementary information S3 in Marchini et al.[43]) but observed no significant biases in null or causal simulations (Tables S8 and S9), most likely as a result of the stringent postimputation QC.

We evaluated multiple other complex architectures with respect to LD (see Appendix A) but observed significant bias in only one deliberately severe scenario: causal variants sampled from intronic and intergenic regions either directly adjacent to or proximal to a DHS (within 1 kb of a DHS boundary). Although no substantial false DHS heritability was observed in genotyped SNPs, the imputed DHS component picked up 50% (0–500 bp) and 20% (500–1,000 bp) of the non-DHS h_g^2 (Figure S11). Given our findings that genotyped SNPs are expected to greatly underestimate DHS enrichment, we consider genotyped and imputed estimates to be lower and upper bounds, respectively, on the true causal enrichment.

Heritability of Functional Categories across 11 Diseases

We analyzed a total of 11 WTCCC1 and WTCCC2 phenotypes.[44–46] After QC,[21] the seven WTCCC1 traits each included an average of 1,700 affected subjects and a set of 2,700 shared control subjects; the four WTCCC2 traits included 1,800–9,300 affected subjects and 5,300 shared control subjects (see Material and Methods; Table S1). In all analyses of autoimmune traits, SNPs in the well-studied MHC region were excluded, although inclusion of the MHC as a separate component did not significantly affect the results. Each cohort was imputed to the 1000 Genomes reference panel, yielding four to six million SNPs per trait after QC (see Material and Methods; Table S1). This analysis is expected to be skewed toward the autoimmune traits, which composed 6/11 traits analyzed and 20,461/30,158 affected subjects analyzed. We computed meta-analysis values by using inverse-variance weighting with the analytical SE to account for different levels of error across h_g^2 estimates. After meta-analysis, resulting SEs were adjusted for the use of shared controls by genomic control (unless otherwise stated), and p values were computed by a simple Z score comparing the mean enrichment and adjusted SE to a null of 1.0 enrichment. Estimating enrichment from shifted functional annotations yielded null enrichments and p values (Tables S10 and S11), confirming that this null is comparable to random SNP comparisons used in previous work.[3,11,40,47]

Combined results meta-analyzed across all traits are reported in Figure 3 (Tables S10, S12, and S13). In genotyped data, DHS variants (spanning 24% of genotyped SNPs) were the most significantly enriched and explained an average of 38% (SE = 4%) of the total h_g^2, a 1.6× enrichment (p = 1.0 × 10^{-4}). Coding variants were the only other category significantly enriched (after six tests were accounted for) and explained 4% (SE = 1%; p = 1.1 × 10^{-3}). All enrichments or depletions were stronger when imputed SNPs were analyzed in terms of both significance and information content, consistent with our previous simulations (Figures S9 and S10; Table S16). Variants in DHSs again exhibited the greatest h_g^2 and most significant enrichment: imputed DHS SNPs explained an average of 79% (SE = 8%) of the total h_g^2, a 5.1× enrichment (p = 3.7 × 10^{-17}). The enrichment varied across traits (Figure S12; Table S14), and there was a nominally significant difference between the six autoimmune traits (AS, CD, MS, RA, T1D, and UC) and the five nonautoimmune traits (SP, bipolar disorder, CAD, hypertension, and type 2 diabetes [T2D]) at 5.5× and 3.3×, respectively (p = 0.01 for difference without accounting for shared control subjects). Coding variants exhibited the greatest overall enrichment at 13.8× (p = 1.8 × 10^{-3}) but accounted for 8% of h_g^2 because of the much smaller category size. Correspondingly, we observed a significant depletion for both intronic regions (0.1×; p = 4.9 × 10^{-9}) and intergenic regions (−0.1×; p < 10^{-20}) and h_g^2 that was not significantly different from 0. We note that compared to genotyped SNPs, imputation in these traits generally does not explain additional h_g^2,[21] but it can more precisely partition heritability into functional categories. We performed additional simulations mimicking the enrichment observed in imputed data with 8,300 causal variants (as inferred in a large GWAS of a polygenic trait[48]) and found that 79% of heritability was explained by imputed DHS SNPs, 8% was explained by imputed coding SNPs, and the remainder was uniformly drawn from the other variant categories. This "realistic" scenario yielded much weaker estimates of enrichment from genotyped SNPs, and they were similar to estimates from genotyped SNPs in real data (Figure 3).

We considered alternative estimation procedures to rule out potential biases. Although we allowed individual values of h_g^2 to fluctuate outside the 0–1 bound on variance to achieve unbiased estimates prior to averaging across traits,[49] a constrained analysis yielded similar results (see Table S15). Individual point estimates escaping the 0–1 bound were consistent with our imputed simulations under realistic enrichment, which showed that the percentage of h_g^2 for DHSs exceeded 1.0 10% of the time, whereas the percentage of h_g^2 for intronic and intergenic regions fell below 0.0 30% and 23% of the time, respectively, for a typical 7,000-sample cohort. Using flat instead of inverse-variance weighting yielded a comparable estimate such that DHS SNPs explained an average of 85% (SE = 15%) of h_g^2. With the flat weighting, the SD of imputed DHS estimates across different traits was 48%, which corresponds to a SD of 32% in the true unobserved values after the analytical SE of each estimate is accounted for (Table S14). We further evaluated the robustness of these estimates and found that biases arising from analytical SEs, ancestry, or case-control ascertainment were unlikely to significantly affect the enrichment (see Appendix A).
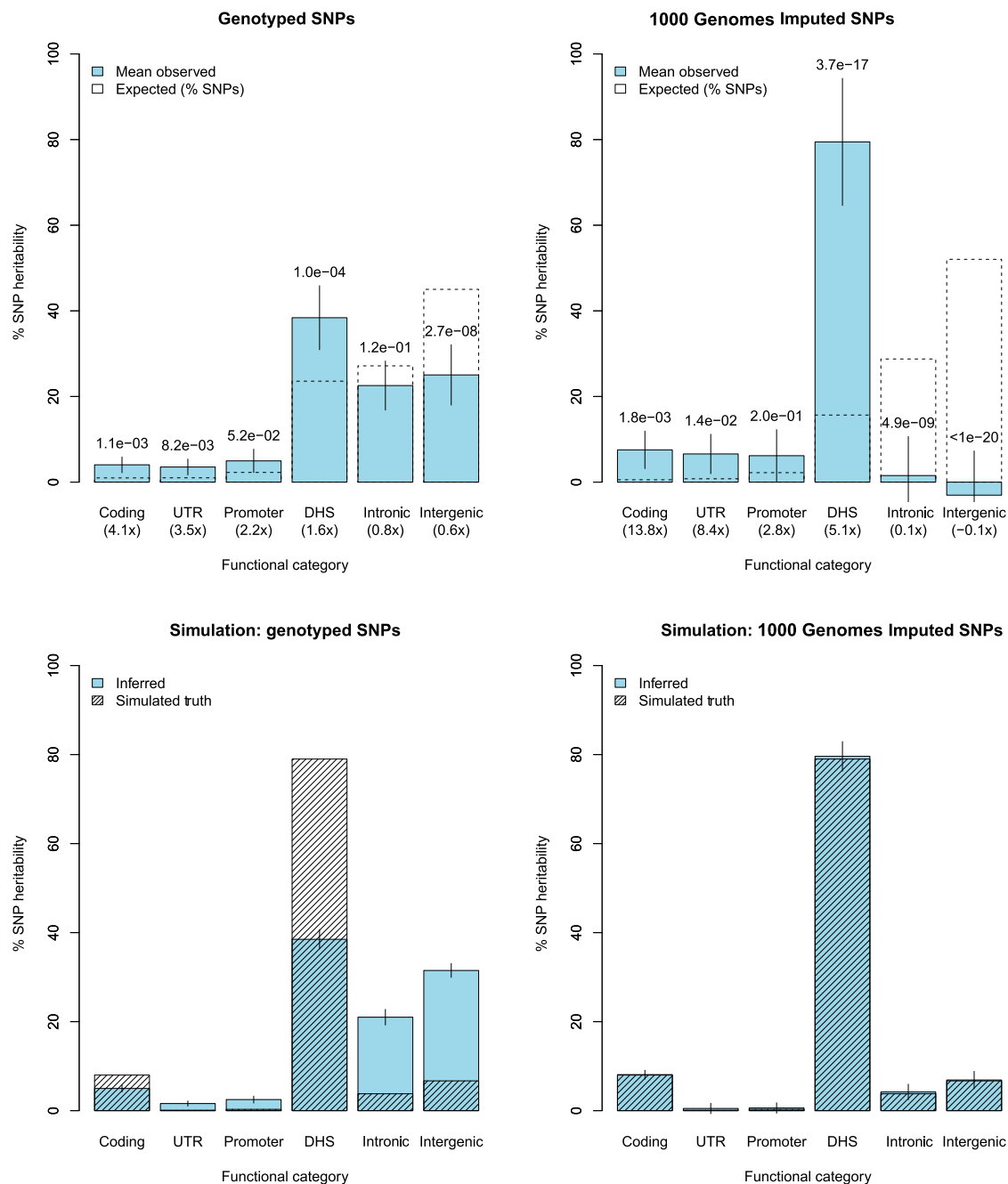
Figure 3. Functional Partitioning of SNP Heritability across 11 Traits

(Top panels) Joint estimates of the percentage of h_g^2 from six functional components are shown in filled bars (meta-analyzed over 11 traits). The null expectation, equal to the percentage of SNPs in each category, is shown by dashed, unfilled bars, and p values report the difference from this expectation. Fold enrichment relative to the null expectation is shown in parentheses below each category. The left panel shows results from analyses of genotyped SNPs only, and the right panel shows analysis of genotyped and 1000 Genomes imputed SNPs. Error bars show 1.96× SE after adjustment for shared controls.

(Bottom panels) Partitioned h_g^2 in simulations of a "realistic" trait where DHS and coding variants explained 79% and 8% of h_g^2, respectively (with no enrichment elsewhere). Filled bars show the mean inferred percentage of h_g^2 from genotyped (left) and imputed (right) SNPs over 100 simulations. Patterned bars show the simulated true partition. Error bars show 1.96× SE (on average, SEs on imputed data were 2.2× higher than SEs on genotype data as a result of the abundance of new variants).

To investigate whether enrichment in h_g^2 from all SNPs at known loci was consistent with the genome-wide estimates, we partitioned the h_g^2 explained by SNPs within 1 Mb of published GWAS loci for each trait (NHGRI GWAS catalog;[11] see Web Resources) (Figure S13). Because some traits had a small number of loci, the DHS component was jointly analyzed with only a single other component containing all non-DHS SNPs. We again observed a highly significant DHS enrichment in imputed data and a significant difference between the genotyped and imputed
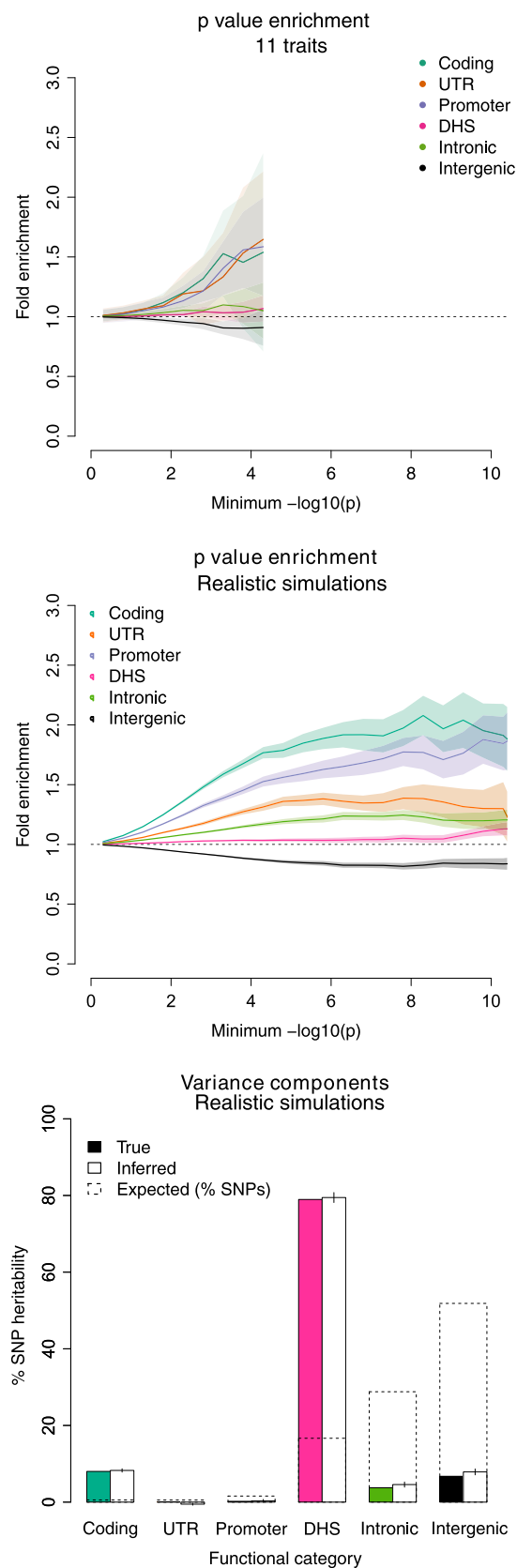
Figure 4. Enrichment from GWAS Summary Statistics
(Left panel) Estimates of p value enrichment are averaged over 11 analyzed traits and are restricted to minimum p value thresholds (x axis) for which at least one association meeting the threshold was observed in every trait.

estimates (p = 7.3 × 10^{-14}). We observed a marginally significant difference between the DHS enrichment at known loci versus genome-wide in the imputed data (3.6× versus 5.5×, p = 0:003). Although it does not pass multiple-test correction, this p value suggests that genome-wide-significant SNPs of large effects might be less enriched with DHS variants than the rest of the genome.

We have shown by simulation that estimates from genotyped SNPs are expected to provide a lower bound on enrichment or depletion and that estimates from imputed SNPs are biased upward only when causal variants are very close to the annotation boundary. For brevity, subsequent results focus primarily on the analysis of imputed SNPs.

Comparison to Estimates of Enrichment from Summary Statistics

We compared our imputed variance-component estimates of 5.1× DHS enrichment for the 11 traits to the DHS enrichment of genome-wide-significant variants identified in these data or from published loci (NHGRI GWAS catalog;[11] see Web Resources). The enrichments from genome-wide-significant variants were much smaller (0.91× and 1.74× for variants in these data and published loci, respectively; Table S17). This is roughly consistent with previous results indicating that 57% of noncoding GWAS hits (from any trait) lie in broadly defined DHSs spanning 42% of the genome (1.4× noncoding enrichment; 1.2× overall enrichment) and that this percentage increases to 77% of noncoding GWAS hits when SNPs in perfect LD with a DHS SNP are included (1.8× noncoding enrichment; 1.6× overall enrichment).[3] Similarly, 30% of the noncoding GWAS hits analyzed in Maurano et al.[3] lay in our DHS annotation, yielding a comparable 1.8× noncoding enrichment. Extending to the full distribution of association statistics did not reveal significant DHS enrichment in any of these traits (Figure 4, left panel; Figure S14). This is consistent with our previous simulations showing the variance-component approach to be more effective than the p-value-enrichment approach at identifying DHS enrichment from complex-disease architectures (Figure 2).

We sought to further confirm this observation by extending our simulations to a single large cohort with realistic levels of enrichment on the basis of the above results. We simulated the "realistic" level of enrichment (see above) in 33,000 combined WTCCC2 samples, corresponding to a large GWAS. We then conducted a standard GWAS on the simulated traits and plotted functional enrichment by using p value enrichment (see Material and Methods). The strategy yielded enrichment at coding

(Middle panel) p value enrichment from a "realistic" simulation. (Right panel) Variance-component enrichment from a "realistic" simulation. Realistic traits were simulated with DHS and coding variants explaining 79% and 8% of h_g^2, respectively, and with computed GWAS statistics in a cohort of 32,000 samples. Shaded regions and error bars represent the SE from meta-analysis (left) and 50 replicates (middle and right).
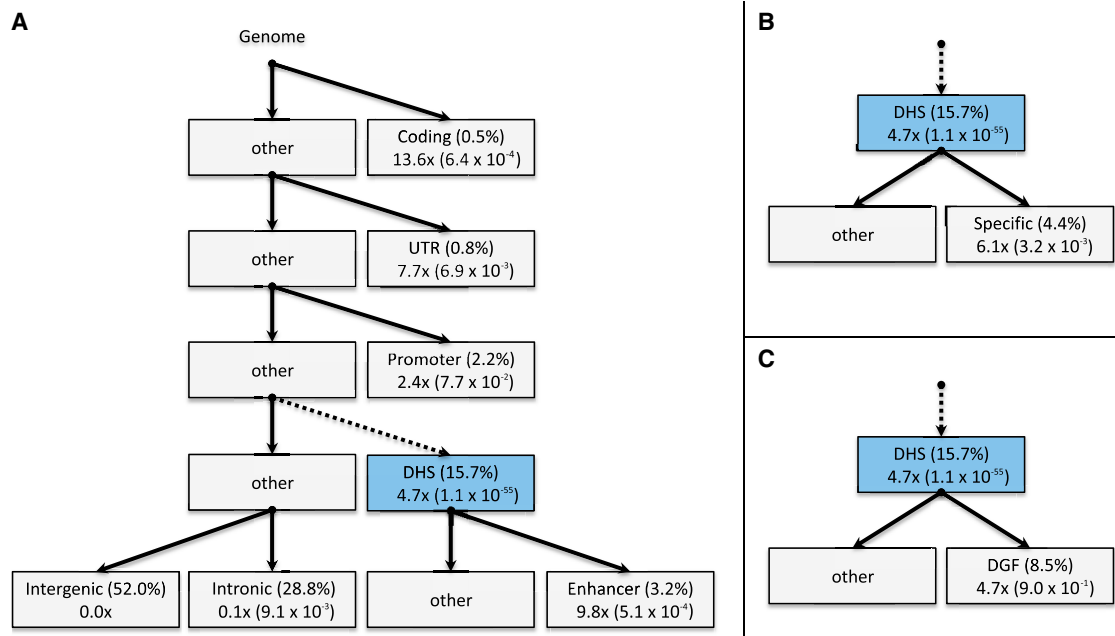
Figure 5. Hierarchical Analysis of Functional Enrichment

DHS variants were further partitioned into three subcategories: predicted enhancers (A), cell-type-specific DHSs (B), and DGF targets (C). Each block contains (on the top line) the functional category and fraction of the genome (in parentheses) and (on the bottom line) the fraction enriched in relation to the rest of the genome and the p value of enrichment in relation to the parent category (in parentheses). DHS enrichment of 4.7× nonsignificantly differed from 5.1× in Figure 3 as a result of additional free parameters.

variants through the full distribution of association statistics (Figure 4, middle panel). However, proximal categories such as UTR and promoter, which were truly depleted, also appeared enriched through tagging of significant coding variants. DHS variants were the least-enriched nonintergenic category, even though they made the single largest contribution to heritability. This was likely due to lower power to detect DHS SNPs as a result of their lower average effect size (relative to that of coding SNPs) and less LD. On the other hand, applying the variance-component strategy to the simulated cohorts correctly recovered the enrichment factors (Figure 4, right panel). These simulations further demonstrate that GWAS p values, although partially informative, can yield false-positive and false-negative enrichment to make functional interpretation difficult, motivating further development of methods that can produce robust estimates of partitioned heritability from summary statistics.

Analysis of PGC2 SP Data

We replicated our functional-enrichment results in an independent cohort of 58,197 samples from PGC2 (Tables S1 and S2). In the PGC2 data, the imputed DHS enrichment was significant at 3.2× (SE = 0.29, p = 1.4 × 10^{-13}), and the intergenic category was significantly depleted at 0.3× (SE = 0.06, depletion p < 1 × 10^{-20}; Table S18). For comparison, the WTCCC2 analysis restricted to SP produced a nonsignificant DHS enrichment of 2.6× (SE = 1.47, p = 0.28) and intergenic h_g^2 of 0.4× (SE = 0.27, depletion p = 0.02; Table S14). The consistency of

WTCCC2 and PGC2 estimates indicates that platform artifacts are unlikely to be a major confounder. Moreover, the substantially lower SE in this large cohort demonstrates the effectiveness of our methods at characterizing a single complex trait. As in our previous simulations, p value enrichment did not identify substantial enrichment at DHS variants (Figure S15).

Partitioning h_g^2 within DHSs

We further partitioned DHS enrichment in the WTCCC1 data into functional subcategories to assess significance in relation to all DHSs. We used Segway-chromHMM combined classifications of enhancer regions[40] to partition DHSs (15.7% of the genome) into those that overlapped predicted enhancers (3.2% of the genome) and those that did not (Figure 5A). The enhancer DHSs explained 31.7% (SE = 3.3%) of the total h_g^2, yielding an enrichment of 9.8× versus all SNPs (1.9× versus all DHSs; p = 5.1 × 10^{-4}). We also partitioned DHSs into regions that were called in two or fewer cell types ("specific"; after merging similar tissues) and those that were not (Figure 5B). We observed a significant enrichment for cell-type-specific DHSs (6.1× versus all SNPs; 1.3× versus all DHSs; p = 3.2 × 10^{-3}). The enrichment was not significant when we repeated this analysis for enhancer and nonenhancer DHSs separately. We next split the DHSs into SNPs overlapping and not overlapping the ENCODE database of DGF regions (8.5% of the genome), which are expected to precisely map sites where regulatory factors bind to the genome[50] (Figure 5C). We observed no difference in h_g^2

Table 1. Cell-Type- and Phenotype-Specific DHS Enrichment

Tissue Type	Cell Type	Autoimmune		Nonautoimmune	
		Genotyped	Imputed	Genotyped	Published
Blood	Primary T helper 1 cells	$5.8 (4.2 \times 10^{-6})$	$10.2 (1.3 \times 10^{-12})$	$2.1 (3.5 \times 10^{-1})$	Maurano et al.[3] (CD)
	leukemia cells	$3.5 (6.7 \times 10^{-6})$	$4.7 (5.3 \times 10^{-10})$	$1.0 (9.8 \times 10^{-1})$	–
	lymphoblastoid cells	$3.3 (1.1 \times 10^{-5})$	$4.9 (5.4 \times 10^{-11})$	$1.0 (9.4 \times 10^{-1})$	Maurano et al.[3] (MS)
	CD8^{+} primary cells	$3.0 (3.0 \times 10^{-4})$	$5.4 (1.8 \times 10^{-10})$	$1.0 (9.6 \times 10^{-1})$	Trynka et al.[6] (RA)
Fetal kidney	fetal right renal pelvic cells	$5.4 (1.4 \times 10^{-4})$	$8.2 (5.7 \times 10^{-8})$	$1.5 (7.4 \times 10^{-1})$	–
Bone marrow	CD14^{+} monocytes	$4.1 (1.6 \times 10^{-4})$	$5.7 (2.2 \times 10^{-7})$	$1.3 (7.6 \times 10^{-1})$	Maurano et al.[3] (MS)
Fetal thymus	Fetal thymus cells	$2.6 (4.0 \times 10^{-4})$	$4.5 (3.2 \times 10^{-9})$	$0.8 (6.6 \times 10^{-1})$	–

Fold enrichment of h_g^2 reported for cell-type-specific DHSs observed as significant in genotype data (after adjustment for 83 cell types tested). We measured enrichment in comparison to h_g^2 at DHSs to account for the background DHS enrichment. Results are shown separately from meta-analyses of six autoimmune traits and five nonautoimmune traits. Instances where enrichment was also observed in Trynka et al.[6] or Maurano et al.[3] are indicated.

between these DHSs and other DHSs (1.0×, p = 0.90). However, DGF annotations were collected for only a subset of DHS cell types analyzed, and analysis in additional cell types is needed. Lastly, we partitioned the h_g^2 by using an expanded DHS annotation (including regions overlapping coding regions, UTRs, and promoters) into the remaining five major categories (Table S19), which yielded 34.4× enrichment at DHS coding variants versus all SNPs (5.3× versus all DHSs, p = 1.35×10^{-3}) and 13.2× enrichment at DHS promoter variants versus all SNPs (2.3× versus all DHSs, p = 7.90×10^{-3}). Notably, unlike the non-DHS introns, DHS introns did not show substantial depletion (0.9× versus all DHSs, p = 0.037).

To investigate the role of specific cell types, we separately estimated enrichment in h_g^2 for DHSs in each of 83 unique cell types (see Material and Methods). For each trait and cell type, we estimated h_g^2 jointly from three components corresponding to DHSs observed in that cell type, other DHSs not observed in that cell type, and all other SNPs; we assessed enrichment in relation to all DHSs. On the basis of our previous observation of heterogeneity, we performed meta-analyses across the six autoimmune traits (excluding the MHC) and across the five nonautoimmune traits. We observed seven cell types that were significantly enriched in autoimmune traits in genotype data (we conservatively adjusted for 83 tests, although the cell types are highly correlated), and none were significantly enriched in nonautoimmune traits (Table 1). Four of these seven cell types have previously been implicated in autoimmune diseases: Trynka et al.[6] found that GWAS hits for RA were enriched within H3K36me3 peaks from CD8^{+} primary cells (at p = 0.0042), and Maurano et al.[3] found that nominally significant SNPs in a GWAS of CD were enriched within DHS peaks from primary T helper 1 cells and that nominally significant SNPs in a GWAS of MS were enriched in DHS peaks from lymphoblastoid and monocyte CD14^{+} cells. The remaining three significant cell types were leukemia cells, fetal pelvis cells, and fetal thymus cells (additional nominally significant cell

types are listed in Table S20). The enrichment was typically observed in all autoimmune traits individually; CD was the least enriched on average (2.8×), and UC was the most enriched on average (5.1×; Table S21). As before, the signal was stronger and more significant when we included imputed SNPs (Table 1).

On the basis of the hypothesis that most regulatory sites lie at the center of the called DHS peaks, we considered the enrichment after progressively narrowing the DHS annotations. Specifically, we trimmed the ends of each DHS peak (without removing any individual peaks) to a maximum length set such that the resulting overall DHS annotation covered 1%, 5%, or 10% of the physical genome. We then tested these three narrowed annotations in two models: (1) a univariate model in which h_g^2 was inferred from only the narrowed DHS component, thereby including any tagged heritability from other functional categories; and (2) a six-component model in which the full DHS component was replaced with the narrowed DHS component and the remaining DHS SNPs were distributed into the intron and other components. We found the DHS centers to be particularly strongly enriched (Table S22); the 1% annotation explained 19.8% of the total h_g^2 in the multivariate model (p = 2.6×10^{-6}) and 61.0% of the total h_g^2 in the univariate model. For comparison, the coding component covering roughly 1% of the genome explained 30.0% of the total h_g^2 in the univariate model. The monotonic increase in h_g^2 from narrowed annotations is further evidence of enrichment at the DHS centers. We caution that this experiment might have been particularly susceptible to bias from causal variants very close to the annotation boundary.

Unbiased Estimates of h_g^2 with Rare and Common Variants

We separately analyzed a cohort of 2,500 SP subjects and 3,875 control subjects who were of homogenous Swedish origin and had been typed on both GWAS and exome chips (see Material and Methods; Tables S1 and S3) to

Table 2. h_g^2 of SP from Exome Chip

Variant Class	h_g^2	Percentage of h_g^2
Separately		
All	0.370 ± 0.040	–
Noncoding	0.317 ± 0.042	–
Coding	0.158 ± 0.034	–
Jointly		
Noncoding	0.291 ± 0.028	79% ± 8%
Coding	0.079 ± 0.034 (p = 1.2 × 10^{-2})	21% ± 6%
Coding (rare)	0.037 ± 0.029 (p = 1.0 × 10^{-1})	10% ± 7%
Coding (common)	0.042 ± 0.017 (p = 7.7 × 10^{-3})	11% ± 4%

Estimates of h_g^2 (adjusted for biases due to LD; see Figure S17 and Table S23) are reported from variance components in the homogenous Swedish subpopulation. The top section shows estimates that include tagging of variants in other classes. The bottom section shows joint estimates accounting for tagged variance due to LD. The p values from a likelihood-ratio test are shown in parentheses.

investigate the possible contribution of rare coding variants to missing heritability,[51] defined as the gap between our genome-wide estimates of h_g^2 and the total narrow-sense heritability. The exome-chip variants were primarily rare and consisted of 18% singletons and 64% nonsingletons with a MAF below 0.01. A concern is that h_g^2 estimates from exome-chip data can be substantially biased as a result of the abundance of rare variants.[21,41,42] To address this, we performed simulations across the full causal-allele frequency spectrum and found that joint estimates from two frequency-stratified[42] components computed from rare (MAF ≤ 0.01) and common (MAF > 0.01) SNPs eliminated most of the observed bias. Subsequently adjusting each component for LD completely eliminated bias for normalized effect sizes (Figure S17) and yielded the most accurate estimate for standard effect sizes (Figure S18). We report estimates from joint components with (Table 2) and without (Table S23) LD adjustment.

We partitioned the heritability explained by GWAS-chip and exome-chip data into three separate variance components: noncoding, rare coding (MAF < 0.01), and common coding variants. This partitioned analysis identified a total h_g^2 of 0.079 (SE = 0.034) from all coding variants (Table 2); only the h_g^2 of 0.042 (SE = 0.017) from common coding variants was significantly different from 0 (p = 7.7 × 10^{-3}; rare coding p = 0:10). Moreover, the estimate of DHS enrichment from common SNPs was unaffected by the inclusion of rare coding variants (Table S24), confirming that DHS enrichment was not an artifact of untagged coding variation in this cohort. The h_g^2 from rare variants remained nonsignificant even after we partitioned according to PolyPhen-2 scores,[52] restriction to putative SP-associated genes (see Appendix A), or gene collapsing (Tables S25 and S41–S43). This does not invalidate the use of collapsed-gene burden tests for association and genetic mapping because the individual collapsed gene is still a fundamentally informative unit of association. It does, however, demonstrate that the maximum variance that can be explained by such methods is guaranteed to be substantially lower than that of association with the full model, as has been shown in previous analyses of burden tests.[53] For singleton variants, we can place a 95% upper bound on collapsed h_g^2 at 0.014. We caution that our exome-chip results pertain to rare variants included in the chip design (ascertained from 12,000 samples) but do not extend to extremely rare variants. However, our findings are consistent with a recent analysis of SP exome sequencing data, which identified a significant but modest rare-coding burden (0.4%–0.6% of total variance) in a subset of ~2,500 genes.[54]

Fine Mapping with Functional Priors

Estimates of functional h_g^2 enrichment can guide fine-mapping analysis, where the goal is to identify a minimal set of SNPs that include the underlying causal variant(s).[55] To investigate the potential benefits of fine mapping on the basis of our estimates of functional enrichment, we applied these estimates as priors for fine mapping in four traits (RA, T2D, CAD, and SP) with publicly available imputed summary statistics (Table S26; see Web Resources). We used corresponding estimates of functional enrichment in the WTCCC1 data for RA, T2D, and CAD (while implicitly assuming a best-case scenario in which functional enrichment was accurately estimated for each trait) and used estimates of functional enrichment in PGC2 data for SP. Given that SNPs at genome-wide-significant loci explain only a small proportion of the trait variance, we do not expect partial sample overlap to be a significant confounder. Although fine-mapping analysis ideally involves targeted sequencing or genotyping, Maller et al.[55] observed that the latter had little impact on their fine-mapping analysis in comparison to imputed data, so we expect imputed markers to be a reasonable proxy. Each locus was defined as the union of 1 Mb windows around any SNP with a p value < 5 × 10^{-8}. Association statistics consisting of individual SNP effect sizes and SEs were converted to Bayes factors as described in Pickrell[13] and Wakefield[56] and were multiplied by either a flat prior or the genome-wide functional prior (computed as the estimated h_g^2 per SNP of the SNP category in the corresponding trait). We then computed the credible set for each locus for each scenario by including SNP Bayes factors from highest to lowest until the sum of the Bayes factors in the set was at least 95% of the sum of the Bayes factors at the locus. On average, we found that the six main functional priors reduced the credible set of causal variants by 30% across the four traits (Table 3). The largest reduction of 63% was observed in RA, where the total credible set for five loci (excluding the MHC) was reduced from 69 SNPs to 26. For comparison, including only coding-variant enrichment as a prior reduced the credible sets by 5% on average and had no reduction for RA. We showed by simulation that the credible sets were well calibrated with the correct priors and miscalibrated by less than 10% when the priors were at

Table 3. Credible Sets of Causal SNPs at Known Associated Loci

Phenotype	No. of Loci	Total SNPs	Flat Prior	Coding Prior	Main Functional Priors	Main and Enhancer Priors
RA	5	8,393	69	69	26	26
T2D	13	24,799	101	90	84	83
CAD	16	27,685	112	112	90	86
CAD (metabo-chip)	34	7,498	325	325	264	260
PGC2	146	582,401	5,696	5,660	4,756	not available

For each trait, genome-wide-significant loci from meta-analysis association statistics were reduced to 95% credible sets with and without functional priors. The right-most four columns describe the number of SNPs in the credible set obtained from each prior type. "Flat prior" corresponds to standard analysis with no functional information. "Coding prior" uses only enrichment at coding variants. "Main functional priors" include all six priors from the main functional analysis. "Main and enhancer priors" include all six main priors and the enhancer-DHS prior.

the extremes of the meta-analysis estimates (Table S27), demonstrating that this functional fine-mapping strategy might become robust and effective as individual trait sample sizes reach the current meta-analysis sample size. However, we caution that our estimates of functional enrichment for individual traits, except SP, are not tight enough for this strategy to be actionable at the current time.

Discussion

The importance of regulatory and cell-type-specific variation in common disease has previously been recognized,[3–10] but in contrast to previous work, we provide a quantification of this contribution to disease heritability. We have demonstrated by extensive simulations that our variance-component strategy yields robust estimates that account for LD between categories and complex-disease architecture. Across 11 traits, we found that regulatory regions marked by DHSs explained an average of 79% of imputed h_g^2 and 38% of genotyped h_g^2. We replicated our results in a large SP cohort, yielding a single-trait estimate of $3.2\times$ (SE = 0.29, p = 1.4×10^{-13}) from imputed SNPs, and found that the contribution from rare, exome-chip variants was nonsignificant and did not affect the enrichment.

Given that GWASs primarily identify noncoding variants, many hypotheses have been developed to explain the architecture of complex traits, including noncoding RNA, DNA methylation, alternative splicing, and unannotated transcripts.[14,57] Several previous studies have demonstrated an excess of significant GWAS associations in regulatory categories.[5,6,11,58] In particular, Ernst et al.[59] observed $2\times$ enrichment in cell-type-relevant enhancers, Schaub et al.[8] identified $1.12\times$ enrichment at DHSs, and Maurano et al.[3] identified $1.4\times - 1.8\times$ enrichment at DHSs (relative to noncoding SNPs) and enrichment at cell-type-relevant DHSs. In our analyzed cohorts, known variants were $1.7\times$ enriched with DHSs, but there was less enrichment at variants identified only in these cohorts. In contrast, our findings constrain most of h_g^2 to the 16% of SNPs that lie in the DHS marks tested (or to SNPs that lie very close to DHSs; see below), particularly in those that overlap enhancers, and suggest that the other pro-

posed mechanisms are unlikely to make substantial independent contributions. A deeper analysis of DHSs narrowed to cover 1% of the genome still explained 20% of h_g^2 directly (and 61% in total), potentially motivating a DHS-targeted genotyping chip analogous to the exome chip.[60] More generally, our approach provides a means of assessing biological hypotheses of contributions to disease heritability.

Unlike previous methods, our approach infers disease-relevant biological function from all SNPs simultaneously instead of one GWAS hit at a time. Over multiple simulated disease architectures, we show that variance-component methods are more accurate in partitioning heritability than summary-statistic-based approaches, such as p value enrichment, despite the appeal of analyses of summary statistics in many contexts.[61–64] For completeness, we also considered two additional methods, stratified Q-Q plots[12] and Bayesian hierarchical modeling (fgwas),[13] which assess functional enrichment but are primarily focused on strong associations and improving mapping power. These methods did not produce consistent estimates of h_g^2 enrichment either in simulations or in real data (Figure S19–S29), although we note that they have different objectives. In addition to having implications for mapping power,[12,13,65–68] functional enrichment has direct implications for fine mapping[55,69,70] and risk prediction. Enrichments at the level we observed could substantially reduce the set of potential causal variants in the four traits we tested by downweighing SNPs in low-heritability categories. On the other hand, the improvement in polygenic risk prediction was limited because of pervasive LD across categories (Table S28).

Several limitations of our approach remain as avenues for future work. The variance-component method might still be subject to subtle biases[21,41,42] under disease architectures or annotations with complex LD structure, although our analyses indicate that it is generally less biased than published methods. In particular, we found that imputed data might lead to an overestimate of category enrichment from causal variation very close to that category. For computational reasons, we did not make use of the mixture of the normal-effect-size approach, which has been shown to increase precision.[24] The method also requires

individual-level genotype data and is computationally infeasible for extremely large cohorts or a very large number of components, motivating further work on methods that analyze summary statistics. A limitation of assessing enrichment from GWAS platforms is that we cannot account for untagged causal variation, which represents roughly half of total narrow-sense heritability.[71] Although we have shown that rare coding variants are unlikely to alter the DHS enrichment, the missing heritability could lie in other categories. The precision of inferred enrichment is also limited by the underlying annotations and variants. It is possible that certain biological features could be subject to systematically poorer variant calling or imputation and exhibit decreased h_g^2 as a result of artifacts,[72] although we did not observe substantial differences in the categories we analyzed. Because of the data available, our meta-analysis estimates were weighted toward autoimmune traits both in the number of individual studies and in total sample size; estimates of DHS enrichment were higher in autoimmune than in nonautoimmune traits, which could be partly due to the abundance of hematopoietic cell types in available DHS annotations. Except for SP, for which many samples are available, we could not provide precise estimates for single traits. However, we have shown by simulation that the individual estimates and errors were well calibrated, justifying meta-analysis of estimates that are not constrained to the plausible 0–1 range (an established strategy[49]). Further partitioning of DHSs can yield additional enrichment, and it is likely that other functional categories—including additional chromatin marks, histone modifications, formaldehyde-assisted isolation of regulatory elements, transcription factor binding sites,[73] gene expression,[58,74,75] and measures of conservation[7]—will be highly informative.

Appendix A

LD

We further interrogated the role of LD and violations of model assumptions in the variance-component estimate. We considered two contrived annotations constructed from either the 16% of SNPs with the most LD partners or the 16% of SNPs with the fewest LD partners to mimic a high or low LD category, respectively, approximately equal in SNP number to the DHS category. Testing the uniformly drawn MAF-independent architecture, we again observed no enrichment for either the high-LD (1.02×, SE = 0.01) or the low-LD (1.02×, SE = 0.03) annotations over 1,000 trials. Finally, we considered a disease architecture in which causal variants were strongly enriched at the centers of DHSs such that variants in the middle 7% of the DHS (1% of the genome) explained 25% of the h_g^2 and the remaining DHS variants explained 75% of the h_g^2. We observed a slight deflation of the DHS estimate, but no significant false enrichment, at the neighboring categories (Figure S20).

Jackknife Estimates of SEs

The analytical SE used for significance testing was accurate in our simulations (Table S29) and has previously been shown to be robust in real data[21,27] but can be biased when the number of causal variants is very small.[41] We assessed this directly with a weighted-block jackknife estimate[76] of the enrichment in the real traits by dropping each chromosome in turn, constructing new GRMs, and recomputing the percentage of h_g^2 for each functional category (and the corresponding enrichment). The jackknife estimate of the enrichment and its variance was then computed as described in Busing et al.[76] Although there is a demonstrable relationship between chromosome length and h_g^2, we do not expect to observe such a relationship with respect to the percentage of h_g^2 because of the normalization. However, this estimate of the variance does capture true biological variation in enrichment across chromosomes and is therefore conservative. Although we observed little difference between the jackknife and standard estimates in genotyped data (Table S30), the jackknife estimate of the imputed percentage of h_g^2 (71%, SE = 7.7%; Table S31) was indeed more conservative than the analytical estimate (79%, SE = 6.6%), but the enrichment was still highly significant (p = 5.5 × 10^{-13}), and the overall results were not substantially affected. Because the jackknife makes no assumptions about the underlying distribution of enrichment, this consistency with the analytical estimate supports the use of REML SEs for case-control data (see also simulations below).

Ancestry

We found little population structure in all of the traits except for MS and SP (Figure S1), which have been previously reported as structured. For the MS cohort, we have shown previously[21] that rigorous ancestry matching did not substantially change the total or partitioned h_g^2. For the SP cohort, we relied on the consistently replicated enrichment across the PGC2 and Swedish SP cohorts, which have been rigorously quality controlled for the avoidance of population stratification. Recently, Janss et al.[77] demonstrated that h_g^2 can vary significantly when principal components are also included as fixed effects as a function of the number of included eigenvectors. To assess the presence of this bias in our Swedish SP data, we recomputed the joint variance-component estimates of $\widehat{h}_g^2$ while including an increasing number of eigenvectors as fixed effects. We observed no significant fluctuation of $\widehat{h}_g^2$ such that the estimates over 1–20 eigenvector covariates had a SD of 0.002, suggesting a tight estimate unbiased by the fixed effects.

Case-Control Ascertainment

Recent work[37,78,79] has shown that liability-scale estimates of h_g^2 from REML can be biased downward in dichotomous traits with strong case-control ascertainment. Golan and Rosset[78] and Hayeck et al.[79] propose an alternative estimator based on Haseman-Elston (H-E) regression[80] and show that it eliminates bias. In brief, this approach

regresses the product of normalized phenotypes on the genetic covariance (off-diagonal GRM entries) for all unique pairs of samples; the resulting slope is used as an estimate of the observed-scale h_g^2 and is converted to the liability scale. This method can be extended naturally to multiple components, where the product of phenotypes is regressed onto GRM entries from each analyzed component in a multiple linear regression. Here, we compared the method and transformation of Golan and Rosset[78] to the REML estimator described in the main text. We also evaluated the impact of incorporating principal components as fixed effects to account for genetic ancestry. This is particularly important for the SP and MS cohorts (see below), which were ascertained in a way that induces correlations between ancestry and phenotype. All analyses were performed with the same set of GRMs computed from 1000 Genomes imputed data, and the H-E regression (and H-E regression with fixed effects) was implemented as described in Golan and Rosset.[78] In all instances, we used analytical error-covariance estimates and rescaled them with the delta method to compute SEs. (We note that the SE for H-E regression makes strongly violated assumptions about independence, and they are therefore only presented for completeness). We observed little difference between variance-component methods and H-E regression methods, and H-E regression yielded an average estimate 1.05× greater than that of REML and an overall $r^2 = 0.95$ between the two methods (across 11 traits; Table S32). The relative performance was similar when we considered only the percentage of h_g^2 from the DHS component (Table S33) such that H-E regression yielded average estimates 1.04× higher than those of REML and an overall $r^2 = 0.94$. When principal components where included as fixed effects, meta-analysis across traits within each method did not yield significant differences (Table S34); H-E regression identified DHS enrichment of 5.8× (SE = 0.45), and REML identified DHS enrichment of 5.1× (SE = 0.42). When we did not include principal components as fixed effects, we observed a large difference between variance components and H-E regression in the SP and MS cohorts, where liability-scale H-E regression estimates of liability-scale h_g^2 were 10.00 and 2.91, respectively (Table S32), outside the plausible 0–1 bound and vastly larger than REML estimates without fixed effects. This suggests that H-E-regression-based estimates might be particularly sensitive to the confounding effects of ancestry.

Lastly, we repeated our null simulations by using the merged WTCCC2 cohort of ~33,000 samples, allowing us to simulate a case-control ascertainment (327 case and 654 control subjects) at a prevalence of 0.01 (see Table S35 for simulation details). When we generated ~1,000 samples on chromosome 1 only, this simulated cohort had an effective SNP-sample ratio (the key quantity driving the effects of case-control ascertainment[37]) corresponding to that of ~10,000 samples genome-wide. We tested a "polygenic" scenario where causal variants were sampled uniformly, as well as a "high-effect" scenario where DHS variants had 10× the effect of other SNPs, and found no significant deviation from the null estimate (Table S35) or the analytical SE (Table S36). Although ascertainment has previously been shown to induce correlation between causal variants, our simulations indicate that this does not bias estimates of enrichment for the prevalence and sample size simulated here.

Detailed Analyses of Rare-Variant h_g^2

Having identified no significant rare-variant h_g^2 at any coding regions, we were interested in quantifying this phenomenon at the set of loci known to be associated with SP. To do so, we constructed six variance components only from SNPs at the 22 loci identified by the PGC1 in a large meta-analysis[48] and estimated h_g^2 jointly with a component for the remaining noncoding variants genome-wide (to account for tagging). As expected, we found the union of all noncoding GWAS variants at these loci to harbor significant heritability of 0.018 (SE = 0.004) (Table S37). However, we did not see any significant heritability from the coding variants at these classes when they were modeled jointly with the other component. This is consistent with our genome-wide finding that common noncoding variants explained a substantial fraction of trait heritability and tagged nearly half of the common coding variation. We also partitioned h_g^2 at the set of 1,796 "composite" genes reported by Purcell et al.[54] to exhibit enrichment of rare disruptive mutations, modeled jointly with exome-chip variants in the remaining genes and noncoding GWAS-chip variants as separate components. However, no significant h_g^2 was observed at either the entire set of composite variants ($h_g^2 = 0.014$, SE = 0.012) or the rare composite variants ($h_g^2 = 0.008$, SE = 0.012).

We observed a significant enrichment in h_g^2 at 4,919 (nonsingleton) loss-of-function variants, which collectively accounted for 6.0% of (nonsingleton) exonic SNPs but explained 24.3% of the exonic h_g^2 (permuted p = 0.02 after MAF matching). We saw no significant enrichment of h_g^2 at coding sites that were predicted to be functionally important by PolyPhen-2.[52] Comparing likelihoods between the model where variants were split into (1) probably damaging and damaging, (2) benign and other, and (3) noncoding components and the model with only (1) coding and (2) noncoding components yielded no significant difference by a 1-degree-of-freedom likelihood-ratio test (p = 0.13).

Supplemental Data

Consortia

The members of the Schizophrenia Working Group of the Psychiatric Genomics Consortium are Stephan Ripke, Benjamin M.

Neale, Aiden Corvin, James T.R. Walters, Kai-How Farh, Peter A. Holmans, Phil Lee, Brendan Bulik-Sullivan, David A. Collier, Hailiang Huang, Tune H. Pers, Ingrid Agartz, Esben Agerbo, Margot Albus, Madeline Alexander, Farooq Amin, Silviu A. Bacanu, Martin Begemann, Richard A. Belliveau, Jr., Judit Bene, Sarah E. Bergen, Elizabeth Bevilacqua, Tim B. Bigdeli, Donald W. Black, Anders D. Børglum, Richard Bruggeman, Nancy G. Buccola, Randy L. Buckner, William Byerley, Wiepke Cahn, Guiqing Cai, Dominique Campion, Rita M. Cantor, Vaughan J. Carr, Noa Carrera, Stanley V. Catts, Kimberly D. Chambert, Raymond C.K. Chan, Ronald Y.L. Chen, Eric Y.H. Chen, Wei Cheng, Eric F.C. Cheung, Siow Ann Chong, C. Robert Cloninger, David Cohen, Nadine Cohen, Paul Cormican, Nick Craddock, James J. Crowley, David Curtis, Michael Davidson, Kenneth L. Davis, Franziska Degenhardt, Jurgen Del Favero, Lynn E. DeLisi, Ditte Demontis, Dimitris Dikeos, Timothy Dinan, Srdjan Djurovic, Gary Donohoe, Elodie Drapeau, Jubao Duan, Frank Dudbridge, Naser Durmishi, Peter Eichhammer, Johan Eriksson, Valentina Escott-Price, Laurent Essioux, Ayman H. Fanous, Martilias S. Farrell, Josef Frank, Lude Franke, Robert Freedman, Nelson B. Freimer, Marion Friedl, Joseph I. Friedman, Menachem Fromer, Giulio Genovese, Lyudmila Georgieva, Elliot S. Gershon, Ina Giegling, Paola Giusti-Rodrguez, Stephanie Godard, Jacqueline I. Goldstein, Vera Golimbet, Srihari Gopal, Jacob Gratten, Jakob Grove, Lieuwe de Haan, Christian Hammer, Marian L. Hamshere, Mark Hansen, Thomas Hansen, Vahram Haroutunian, Annette M. Hartmann, Frans A. Henskens, Stefan Herms, Joel N. Hirschhorn, Per Hoffmann, Andrea Hofman, Mads V. Hollegaard, David M. Hougaard, Masashi Ikeda, Inge Joa, Antonio Julià, René S. Kahn, Luba Kalaydjieva, Sena Karachanak-Yankova, Juha Karjalainen, David Kavanagh, Matthew C. Keller, Brian J. Kelly, James L. Kennedy, Andrey Khrunin, Yunjung Kim, Janis Klovins, James A. Knowles, Bettina Konte, Vaidutis Kucinskas, Zita Ausrele Kucinskiene, Hana Kuzelova-Ptackova, Anna K. Kähler, Claudine Laurent, Jimmy Lee Chee Keong, S. Hong Lee, Sophie E. Legge, Bernard Lerer, Miaoxin Li, Tao Li, Kung-Yee Liang, Jeffrey Lieberman, Svetlana Limborska, Carmel M. Loughland, Jan Lubinski, Jouko Lnnqvist, Milan Macek, Jr., Patrik K.E. Magnusson, Brion S. Maher, Wolfgang Maier, Jacques Mallet, Sara Marsal, Manuel Mattheisen, Morten Mattingsdal, Robert W. McCarley, Colm McDonald, Andrew M. McIntosh, Sandra Meier, Carin J. Meijer, Bela Melegh, Ingrid Melle, Raquelle I. Mesholam-Gately, Andres Metspalu, Patricia T. Michie, Lili Milani, Vihra Milanova, Younes Mokrab, Derek W. Morris, Ole Mors, Preben B. Mortensen, Kieran C. Murphy, Robin M. Murray, Inez Myin-Germeys, Bertram Mller-Myhsok, Mari Nelis, Igor Nenadic, Deborah A. Nertney, Gerald Nestadt, Kristin K. Nicodemus, Liene Nikitina-Zake, Laura Nisenbaum, Annelie Nordin, Eadbhard O'Callaghan, Colm O'Dushlaine, F. Anthony O'Neill, Sang-Yun Oh, Ann Olincy, Line Olsen, Jim Van Os, Psychosis Endophenotypes International Consortium, Christos Pantelis, George N. Papadimitriou, Sergi Papiol, Elena Parkhomenko, Michele T. Pato, Tiina Paunio, Milica Pejovic-Milovancevic, Diana O. Perkins, Olli Pietilinen, Jonathan Pimm, Andrew J. Pocklington, John Powell, Alkes Price, Ann E. Pulver, Shaun M. Purcell, Digby Quested, Henrik B. Rasmussen, Abraham Reichenberg, Mark A. Reimers, Alexander L. Richards, Joshua L. Roffman, Panos Roussos, Douglas M. Ruderfer, Veikko Salomaa, Alan R. Sanders, Ulrich Schall, Christian R. Schubert, Thomas G. Schulze, Sibylle G. Schwab, Edward M. Scolnick, Rodney J. Scott, Larry J. Seidman, Jianxin Shi, Engilbert Sigurdsson, Teimuraz Silagadze, Jeremy M. Silverman, Kang Sim, Petr Slominsky, Jordan W. Smoller, Hon-Cheong So, Chris C.A. Spencer, Eli A. Stahl, Hreinn Stefansson, Stacy Steinberg, Elisabeth Stogmann, Richard E. Straub, Eric Strengman, Jana Strohmaier, T. Scott Stroup, Mythily Subramaniam, Jaana Suvisaari, Dragan M. Svrakic, Jin P. Szatkiewicz, Erik Sderman, Srinivas Thirumalai, Draga Toncheva, Paul A. Tooney, Sarah Tosato, Juha Veijola, John Waddington, Dermot Walsh, Dai Wang, Qiang Wang, Bradley T. Webb, Mark Weiser, Dieter B. Wildenauer, Nigel M. Williams, Stephanie Williams, Stephanie H. Witt, Aaron R. Wolen, Emily H.M. Wong, Brandon K. Wormley, Jing Qin Wu, Hualin Simon Xi, Clement C. Zai, Xuebin Zheng, Fritz Zimprich, Naomi R. Wray, Kari Stefansson, Peter M. Visscher, Wellcome Trust Case Control Consortium, Rolf Adolfsson, Ole A. Andreassen, Douglas H.R. Blackwood, Elvira Bramon, Joseph D. Buxbaum, Anders D. Brglum, Sven Cichon, Ariel Darvasi, Enrico Domenici, Hannelore Ehrenreich, Tõnu Esko, Pablo V. Gejman, Michael Gill, Hugh Gurling, Christina M. Hultman, Nakao Iwata, Assen V. Jablensky, Erik G. Jönsson, Kenneth S. Kendler, George Kirov, Jo Knight, Todd Lencz, Douglas F. Levinson, Qingqin S. Li, Jianjun Liu, Anil K. Malhotra, Steven A. McCarroll, Andrew McQuillin, Jennifer L. Moran, Preben B. Mortensen, Bryan J. Mowry, Markus M. Nthen, Roel A. Ophoff, Michael J. Owen, Aarno Palotie, Carlos N. Pato, Tracey L. Petryshen, Danielle Posthuma, Marcella Rietschel, Brien P. Riley, Dan Rujescu, Pak C. Sham, Pamela Sklar, David St. Clair, Daniel R. Weinberger, Jens R. Wendland, Thomas Werge, Mark J. Daly, Patrick F. Sullivan, and Michael C. O'Donovan.

The members of the SWE-SCZ Consortium are Stephan Ripke, Colm O'Dushlaine, Kimberly Chambert, Jennifer L. Moran, Anna K. Kähler, Susanne Akterin, Sarah Bergen, Patrik K.E. Magnusson, Benjamin M. Neale, Douglas Ruderfer, Edward Scolnick, Shaun Purcell, Steve McCarroll, Pamela Sklar, Christina M Hultman, and Patrick F. Sullivan.

Acknowledgments

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes Phase 1 reference panels, https://mathgen.stats.ox.ac.uk/impute/impute_v2.html#reference

CARDIoGRAM CAD summary statistics, http://www.cardiogramplusc4d.org/downloads/

DIAGRAM T2D summary statistics, http://diagram-consortium.org/downloads.html

DNaseI Digital Genomic Footprinting (DGF) annotations, http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeUwDgf/

Exome Chip Design, http://genome.sph.umich.edu/wiki/Exome_Chip_Design

fgwas, https://github.com/joepickrell/fgwas

Functional annotations, http://www.hsph.harvard.edu/alkes-price/software/

Genome-wide Complex Trait Analysis (GCTA), http://www.complextraitgenomics.com/software/gcta/

HAPI-UR, http://genetics.med.harvard.edu/reich/Reich_Lab/Software.html

IMPUTE2, https://mathgen.stats.ox.ac.uk/impute/impute_v2.html

NHGRI GWAS catalog, http://www.genome.gov/gwastudies/

Oxford recombination map, http://hapmap.ncbi.nlm.nih.gov/downloads/recombination/

Psychiatric Genomic Consortium, Sweden+SCZ1 schizophrenia summary statistics, http://www.med.unc.edu/pgc/downloads

RA summary statistics, http://www.broadinstitute.org/ftp/pub/rheumatoid_arthritis/Stahl_etal_2010NG/

Segway-chromHMM combined enhancer annotations, ftp://ftp.ebi.ac.uk/pub/databases/ensembl/encode/integration_data_jan2011/byDataType/segmentations/jan2011

References

1. Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R., et al. (2010). The nih roadmap epigenomics mapping consortium. Nat. Biotechnol. *28*, 1045–1048.

2. Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M.; ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

3. Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., et al. (2012). Systematic localization of common disease-associated variation in regulatory DNA. Science *337*, 1190–1195.

4. Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75–82.

5. Degner, J.F., Pai, A.A., Pique-Regi, R., Veyrieras, J.-B., Gaffney, D.J., Pickrell, J.K., De Leon, S., Michelini, K., Lewellen, N., Crawford, G.E., et al. (2012). DNase I sensitivity QTLs are a major determinant of human expression variation. Nature *482*, 390–394.

6. Trynka, G., Sandor, C., Han, B., Xu, H., Stranger, B.E., Liu, X.S., and Raychaudhuri, S. (2013). Chromatin marks identify critical cell types for fine mapping complex trait variants. Nat. Genet. *45*, 124–130.

7. Ward, L.D., and Kellis, M. (2012). Evidence of abundant purifying selection in humans for recently acquired regulatory functions. Science *337*, 1675–1678.

8. Schaub, M.A., Boyle, A.P., Kundaje, A., Batzoglou, S., and Snyder, M. (2012). Linking disease associations with regulatory information in the human genome. Genome Res. *22*, 1748–1759.

9. Stergachis, A.B., Haugen, E., Shafer, A., Fu, W., Vernot, B., Reynolds, A., Raubitschek, A., Ziegler, S., LeProust, E.M., Akey, J.M., and Stamatoyannopoulos, J.A. (2013). Exonic transcription factor binding directs codon choice and affects protein evolution. Science *342*, 1367–1372.

10. McVicker, G., van de Geijn, B., Degner, J.F., Cain, C.E., Banovich, N.E., Raj, A., Lewellen, N., Myrthil, M., Gilad, Y., and Pritchard, J.K. (2013). Identification of genetic variants that affect histone modifications in human cells. Science *342*, 747–749.

11. Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., and Manolio, T.A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc. Natl. Acad. Sci. USA *106*, 9362–9367.

12. Schork, A.J., Thompson, W.K., Pham, P., Torkamani, A., Roddey, J.C., Sullivan, P.F., Kelsoe, J.R., O'Donovan, M.C., Furberg, H., Schork, N.J., et al.; Tobacco and Genetics Consortium; Bipolar Disorder Psychiatric Genomics Consortium; Schizophrenia Psychiatric Genomics Consortium (2013). All SNPs are not created equal: genome-wide association studies reveal a consistent pattern of enrichment among functionally annotated SNPs. PLoS Genet. *9*, e1003449.

13. Pickrell, J.K. (2014). Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. Am. J. Hum. Genet. *94*, 559–573.

14. Mudge, J.M., Frankish, A., and Harrow, J. (2013). Functional transcriptomics in the post-ENCODE era. Genome Res. *23*, 1961–1973.

15. Kellis, M., Wold, B., Snyder, M.P., Bernstein, B.E., Kundaje, A., Marinov, G.K., Ward, L.D., Birney, E., Crawford, G.E., Dekker, J., et al. (2014). Defining functional DNA elements in the human genome. Proc. Natl. Acad. Sci. USA *111*, 6131–6138.

16. Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. *42*, 565–569.

17. Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. *88*, 76–82.

18. Yang, J., Manolio, T.A., Pasquale, L.R., Boerwinkle, E., Caporaso, N., Cunningham, J.M., de Andrade, M., Feenstra, B., Feingold, E., Hayes, M.G., et al. (2011). Genome partitioning of genetic variation for complex traits using common SNPs. Nat. Genet. *43*, 519–525.

19. Lee, S.H., Wray, N.R., Goddard, M.E., and Visscher, P.M. (2011). Estimating missing heritability for disease from genome-wide association studies. Am. J. Hum. Genet. *88*, 294–305.

20. Lee, S.H., DeCandia, T.R., Ripke, S., Yang, J., Sullivan, P.F., Goddard, M.E., Keller, M.C., Visscher, P.M., and Wray, N.R.; Schizophrenia Psychiatric Genome-Wide Association Study Consortium (PGC-SCZ); International Schizophrenia Consortium (ISC); Molecular Genetics of Schizophrenia Collaboration (MGS) (2012). Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. Nat. Genet. *44*, 247–250.

21. Gusev, A., Bhatia, G., Zaitlen, N., Vilhjalmsson, B.J., Diogo, D., Stahl, E.A., Gregersen, P.K., Worthington, J., Klareskog, L.,

Raychaudhuri, S., et al. (2013). Quantifying missing heritability at known GWAS loci. PLoS Genet. *9*, e1003993.

22. Lee, S.H., Ripke, S., Neale, B.M., Faraone, S.V., Purcell, S.M., Perlis, R.H., Mowry, B.J., Thapar, A., Goddard, M.E., Witte, J.S., et al.; Cross-Disorder Group of the Psychiatric Genomics Consortium; International Inflammatory Bowel Disease Genetics Consortium (IIBDGC) (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat. Genet. *45*, 984–994.

23. Falconer, D.S. (1965). The inheritance of liability to certain diseases, estimated from the incidence among relatives. Ann. Hum. Genet. *29*, 51–76.

24. Zhou, X., Carbonetto, P., and Stephens, M. (2013). Polygenic modeling with bayesian sparse linear mixed models. PLoS Genet. *9*, e1003264.

25. Fischer, T.M., Gilmour, A.R., and van der Werf, J.H. (2004). Computing approximate standard errors for genetic parameters derived from random regression models fitted by average information REML. Genet. Sel. Evol. *36*, 363–369.

26. Lynch, M., and Walsh, B. (1998). Genetics and Analysis of Quantitative Traits (Sunderland: Sinauer Associates).

27. Visscher, P.M., Hemani, G., Vinkhuyzen, A.A.E., Chen, G.-B., Lee, S.H., Wray, N.R., Goddard, M.E., and Yang, J. (2014). Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. PLoS Genet. *10*, e1004269.

28. Nakaoka, H., and Inoue, I. (2009). Meta-analysis of genetic association studies: methodologies, between-study heterogeneity and winner's curse. J. Hum. Genet. *54*, 615–623.

29. Clayton, D.G., Walker, N.M., Smyth, D.J., Pask, R., Cooper, J.D., Maier, L.M., Smink, L.J., Lam, A.C., Ovington, N.R., Stevens, H.E., et al. (2005). Population structure, differential bias and genomic control in a large-scale, case-control association study. Nat. Genet. *37*, 1243–1246.

30. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909.

31. Williams, A.L., Patterson, N., Glessner, J., Hakonarson, H., and Reich, D. (2012). Phasing of many thousands of genotyped samples. Am. J. Hum. Genet. *91*, 238–251.

32. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J., and Abecasis, G.R. (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat. Genet. *44*, 955–959.

33. Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A.; 1000 Genomes Project Consortium (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56–65.

34. Myers, S., Bottolo, L., Freeman, C., McVean, G., and Donnelly, P. (2005). A fine-scale map of recombination rates and hotspots across the human genome. Science *310*, 321–324.

35. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421–427.

36. Bulik-Sullivan, B., Loh, P.-R., Finucane, H., Ripke, S., Yang, J., Patterson, N., Daly, M.J., Price, A.L., and Neale, B.M.; Schizophrenia Working Group Psychiatric Genomics Consortium (2014). LD score regression distinguishes confounding from polygenicity in genome-wide association studies. bioRxiv http://dx.doi.org/10.1101/002931.

37. Yang, J., Zaitlen, N.A., Goddard, M.E., Visscher, P.M., and Price, A.L. (2014). Advantages and pitfalls in the application of mixed-model association methods. Nat. Genet. *46*, 100–106.

38. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

39. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

40. Hoffman, M.M., Ernst, J., Wilder, S.P., Kundaje, A., Harris, R.S., Libbrecht, M., Giardine, B., Ellenbogen, P.M., Bilmes, J.A., Birney, E., et al. (2013). Integrative annotation of chromatin elements from ENCODE data. Nucleic Acids Res. *41*, 827–841.

41. Speed, D., Hemani, G., Johnson, M.R., and Balding, D.J. (2012). Improved heritability estimation from genome-wide SNPs. Am. J. Hum. Genet. *91*, 1011–1021.

42. Lee, S.H., Yang, J., Chen, G.-B., Ripke, S., Stahl, E.A., Hultman, C.M., Sklar, P., Visscher, P.M., Sullivan, P.F., Goddard, M.E., and Wray, N.R. (2013). Estimation of SNP heritability from dense genotype data. Am. J. Hum. Genet. *93*, 1151–1155.

43. Marchini, J., and Howie, B. (2010). Genotype imputation for genome-wide association studies. Nat. Rev. Genet. *11*, 499–511.

44. Wellcome Trust Case Control Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature *447*, 661–678.

45. Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C.C., Patsopoulos, N.A., Moutsianas, L., Dilthey, A., Su, Z., Freeman, C., Hunt, S.E., et al.; International Multiple Sclerosis Genetics Consortium; Wellcome Trust Case Control Consortium 2 (2011). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature *476*, 214–219.

46. Barrett, J.C., Lee, J.C., Lees, C.W., Prescott, N.J., Anderson, C.A., Phillips, A., Wesley, E., Parnell, K., Zhang, H., Drummond, H., et al.; UK IBD Genetics Consortium; Wellcome Trust Case Control Consortium 2 (2009). Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. Nat. Genet. *41*, 1330–1334.

47. Karczewski, K.J., Dudley, J.T., Kukurba, K.R., Chen, R., Butte, A.J., Montgomery, S.B., and Snyder, M. (2013). Systematic functional regulatory assessment of disease-associated variants. Proc. Natl. Acad. Sci. USA *110*, 9607–9612.

48. Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J.L., Kähler, A.K., Akterin, S., Bergen, S.E., Collins, A.L., Crowley, J.J., Fromer, M., et al.; Multicenter Genetic Studies of Schizophrenia Consortium; Psychosis Endophenotypes International Consortium; Wellcome Trust Case Control Consortium 2 (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat. Genet. *45*, 1150–1159.

49. Wright, F.A., Sullivan, P.F., Brooks, A.I., Zou, F., Sun, W., Xia, K., Madar, V., Jansen, R., Chung, W., Zhou, Y.-H., et al. (2014). Heritability and genomics of gene expression in peripheral blood. Nat. Genet. *46*, 430–437.

50. Neph, S., Vierstra, J., Stergachis, A.B., Reynolds, A.P., Haugen, E., Vernot, B., Thurman, R.E., John, S., Sandstrom, R., Johnson, A.K., et al. (2012). An expansive human regulatory lexicon encoded in transcription factor footprints. Nature *489*, 83–90.

51. Gibson, G. (2011). Rare and common variants: twenty arguments. Nat. Rev. Genet. *13*, 135–145.

52. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat. Methods *7*, 248–249.

53. Liu, D.J., and Leal, S.M. (2012). Estimating genetic effects and quantifying missing heritability explained by identified rare-variant associations. Am. J. Hum. Genet. *91*, 585–596.

54. Purcell, S.M., Moran, J.L., Fromer, M., Ruderfer, D., Solovieff, N., Roussos, P., O'Dushlaine, C., Chambert, K., Bergen, S.E., Kähler, A., et al. (2014). A polygenic burden of rare disruptive mutations in schizophrenia. Nature *506*, 185–190.

55. Maller, J.B., McVean, G., Byrnes, J., Vukcevic, D., Palin, K., Su, Z., Howson, J.M.M., Auton, A., Myers, S., Morris, A., et al.; Wellcome Trust Case Control Consortium (2012). Bayesian refinement of association signals for 14 loci in 3 common diseases. Nat. Genet. *44*, 1294–1301.

56. Wakefield, J. (2009). Bayes factors for genome-wide association studies: comparison with P-values. Genet. Epidemiol. *33*, 79–86.

57. Edwards, S.L., Beesley, J., French, J.D., and Dunning, A.M. (2013). Beyond GWASs: illuminating the dark road from association to function. Am. J. Hum. Genet. *93*, 779–797.

58. Nicolae, D.L., Gamazon, E., Zhang, W., Duan, S., Dolan, M.E., and Cox, N.J. (2010). Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. PLoS Genet. *6*, e1000888.

59. Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43–49.

60. Huyghe, J.R., Jackson, A.U., Fogarty, M.P., Buchkovich, M.L., Stančáková, A., Stringham, H.M., Sim, X., Yang, L., Fuchsberger, C., Cederberg, H., et al. (2013). Exome array analysis identifies new loci and low-frequency variants influencing insulin processing and secretion. Nat. Genet. *45*, 197–201.

61. Lee, S., Teslovich, T.M., Boehnke, M., and Lin, X. (2013). General framework for meta-analysis of rare variants in sequencing association studies. Am. J. Hum. Genet. *93*, 42–53.

62. Lee, D., Bigdeli, T.B., Riley, B.P., Fanous, A.H., and Bacanu, S.-A. (2013). DIST: direct imputation of summary statistics for unmeasured SNPs. Bioinformatics *29*, 2925–2927.

63. Pasaniuc, B., Zaitlen, N., Shi, H., Bhatia, G., Gusev, A., Pickrell, J., Hirschhorn, J., Strachan, D.P., Patterson, N., and Price, A.L. (2014). Fast and accurate imputation of summary statistics enhances evidence of functional enrichment. Bioinformatics *30*, 2906–2914.

64. Liu, D.J., Peloso, G.M., Zhan, X., Holmen, O.L., Zawistowski, M., Feng, S., Nikpay, M., Auer, P.L., Goel, A., Zhang, H., et al. (2014). Meta-analysis of gene-level tests for rare variant association. Nat. Genet. *46*, 200–204.

65. Sun, L., Craiu, R.V., Paterson, A.D., and Bull, S.B. (2006). Stratified false discovery control for large-scale hypothesis testing with application to genome-wide association studies. Genet. Epidemiol. *30*, 519–530.

66. Eskin, E. (2008). Increasing power in association studies by using linkage disequilibrium structure and molecular function as prior information. Genome Res. *18*, 653–660.

67. Darnell, G., Duong, D., Han, B., and Eskin, E. (2012). Incorporating prior information into association studies. Bioinformatics *28*, i147–i153.

68. Carbonetto, P., and Stephens, M. (2013). Integrated enrichment analysis of variants and pathways in genome-wide association studies indicates central role for IL-2 signaling genes in type 1 diabetes, and cytokine signaling genes in Crohn's disease. PLoS Genet. *9*, e1003770.

69. Gaffney, D.J., Veyrieras, J.-B., Degner, J.F., Pique-Regi, R., Pai, A.A., Crawford, G.E., Stephens, M., Gilad, Y., and Pritchard, J.K. (2012). Dissecting the regulatory architecture of gene expression QTLs. Genome Biol. *13*, R7.

70. Gagliano, S.A., Barnes, M.R., Weale, M.E., and Knight, J. (2014). A Bayesian method to incorporate hundreds of functional characteristics with association evidence to improve variant prioritization. PLoS ONE *9*, e98122.

71. Visscher, P.M., Brown, M.A., McCarthy, M.I., and Yang, J. (2012). Five years of GWAS discovery. Am. J. Hum. Genet. *90*, 7–24.

72. Green, P., and Ewing, B. (2013). Comment on "Evidence of abundant purifying selection in humans for recently acquired regulatory functions". Science *340*, 682.

73. Pique-Regi, R., Degner, J.F., Pai, A.A., Gaffney, D.J., Gilad, Y., and Pritchard, J.K. (2011). Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. Genome Res. *21*, 447–455.

74. Lappalainen, T., Sammeth, M., Friedländer, M.R., 't Hoen, P.A., Monlong, J., Rivas, M.A., Gonzàlez-Porta, M., Kurbatova, N., Griebel, T., Ferreira, P.G., et al.; Geuvadis Consortium (2013). Transcriptome and genome sequencing uncovers functional variation in humans. Nature *501*, 506–511.

75. Davis, L.K., Yu, D., Keenan, C.L., Gamazon, E.R., Konkashbaev, A.I., Derks, E.M., Neale, B.M., Yang, J., Lee, S.H., Evans, P., et al. (2013). Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. PLoS Genet. *9*, e1003864.

76. Busing, F.M., Meijer, E., and Van Der Leeden, R. (1999). Delete-m jackknife for unequal m. Stat. Comput. *9*, 3–8.

77. Janss, L., de Los Campos, G., Sheehan, N., and Sorensen, D. (2012). Inferences from genomic models in stratified populations. Genetics *192*, 693–704.

78. Golan, D. and Rosset, S. (2013). Narrowing the gap on heritability of common disease by direct estimation in case-control GWAS. arXiv, arXiv:1305.5363, http://arxiv.org/abs/1305.5363.

79. Hayeck, T., Zaitlen, N., Loh, P.-R., Vilhjalmsson, B., Pollack, S., Gusev, A., Yang, J., Chen, G.-B., Goddard, M.E., Visscher, P.M., et al. (2014). Mixed model with correction for case-control ascertainment increases association power. bioRxiv http://dx.doi.org/10.1101/008755.

80. Haseman, J.K., and Elston, R.C. (1972). The investigation of linkage between a quantitative trait and a marker locus. Behav. Genet. *2*, 3–19.