Genome-wide fine-mapping improves identification of causal variants

Yang Wu[1,2,*], Zhili Zheng[2,3,4], Loic Thibaut[2], Michael E. Goddard[5,6], Naomi R. Wray[2,7], Peter M. Visscher[2,8], Jian Zeng[2,*]

[1] Institute of Rare Diseases, West China Hospital of Sichuan University, Chengdu, China

[2] Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia

[3] Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA

[4] Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA

[5] Faculty of Veterinary and Agricultural Science, University of Melbourne, Parkville, Victoria, Australia

[6] Biosciences Research Division, Department of Economic Development, Jobs, Transport and Resources, Bundoora, Victoria, Australia

[7] Department of Psychiatry, University of Oxford, Oxford, UK

[8] Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, UK

[*] Correspondence: Yang Wu (yang.wu@wchscu.edu.cn), Jian Zeng (j.zeng@uq.edu.au)

Abstract

Fine-mapping refines genotype-phenotype association signals to identify causal variants underlying complex traits. However, current methods typically focus on individual genomic segments without considering the global genetic architecture. Here, we demonstrate the advantages of performing genome-wide fine-mapping (GWFM) and develop methods to facilitate GWFM. In simulations and real data analyses, GWFM outperforms current methods in error control, mapping power and precision, replication rate, and trans-ancestry phenotype prediction. For 48 well-powered traits in the UK Biobank, we identify causal variants that collectively explain 17% of the SNP-based heritability, and predict that fine-mapping 50% of that would require 2 million samples on average. We pinpoint a known causal variant, as proof-of-principle, at FTO for body mass index, unveil a hidden secondary variant with evolutionary conservation, and identify new missense causal variants for schizophrenia and Crohn's disease. Overall, we analyse 599 complex traits with 13 million SNPs, highlighting the efficacy of GWFM with functional annotations.

Introduction

Genome-wide association studies (GWAS) have successfully identified numerous genetic variants associated with complex traits[1-3]. However, the underlying casual variants for these traits are largely unknown. In a standard GWAS, the trait-variant associations are tested one at a time, leading to the discovery of clusters of mutually correlated marginal-association signals due to linkage disequilibrium (LD) between SNPs[4]. While post-GWAS methods such as LD clumping[5] or COJO[6] are used to identify independently significant association signals, SNPs prioritised by these methods are not necessarily the causal variants[7,8].

Statistical fine-mapping, often employing a Bayesian mixture model (BMM), offers a direct approach to narrow down the likely causal variants[9]. In contrast to GWAS, which test marginal effects, fine-mapping aims to detect joint-association signals for causal inference, where the strength of joint association is assessed using the posterior inclusion probability (PIP). PIP is the probability of a SNP being included with a nonzero effect in the model, which, in theory, controls false discovery rate (FDR)[10]. Due to the computational burden and complexity of test hypotheses, current fine-mapping methods focus on genome-wide significant loci only or consider one genomic region at a time (e.g., a LD block), in isolation from the rest of the genome[11-14]. Methods differ mainly in the algorithm used to derive PIP. For example, FINEMAP[12] utilizes a shotgun stochastic search algorithm to explore possible causal configurations, and computes the PIP by averaging over those with non-negligible probabilities. SuSiE[11] and SuSiE-RSS[14] assume a sparse effect model and employ an iterative Bayesian stepwise selection approach to estimate the overall effect of each SNP by summing up multiple singe-effect vectors. SuSiE-Inf[13] and FINEMAP-Inf[13] further extend the two models to include an infinitesimal component for improved modelling of polygenic architecture within a locus.

Despite being widely used, region-specific analysis has limitations. First, the prior specification of genetic architecture is crucial, but is often conservatively predetermined in these analysis[11,12,15] (e.g., prior probability of association set to one over the number of SNPs in the region), which can result in reduced power. Second, fine-mapping can benefit from incorporating functional genomic annotations[16-18], but region-specific methods require step-wise procedures so that GWAS data and functional annotations are not modelled jointly[19]. Third, none of the current methods estimates the power of identifying the causal variants for a trait, which is critical to inform the

experimental design of a prospective study (such a power analysis is available in GWAS[20] but not in fine-mapping).

These limitations of current fine-mapping methods can be addressed through conducting a fine-mapping analysis using a genome-wide Bayesian mixture model (GBMM). GBMMs, which have been widely used for predicting breeding values in agricultural species[21-23] and complex trait phenotypes in humans[24-27], have recently emerged as a method of GWFM[28,29]. Compared to conventional GWAS and region-specific fine-mapping approaches, GBMMs consider genome-wide SNPs simultaneously, which are all utilised to estimate the genetic architecture and functional prior[27,28]. For example, SNPs with the same class of functional annotation are present across the genome. By considering all SNPs jointly, the importance of a functional annotation in a local genomic region can be prioritised based on the evidence for association as a class across the genome. In GBMMs, Markov chain Monte Carlo (MCMC) sampling is often used for posterior inference, which is asymptotically exact and superior to the variational inference regarding accuracy[30], but computationally challenging when analysing high-density SNPs. Fortunately, recent advances in methodology, such as SBayesRC[27], have allowed fitting all common SNPs efficiently in a MCMC-based GBMM. Moreover, GBMMs estimate the polygenicity and variant effect size distribution[22,24,26,27,31,32], providing an opportunity to predict the power of prospective studies with larger sample sizes. However, relevant theory and methods have not yet been developed.

In this study, we comprehensively assess the performance of GWFM analysis using a GBMM (**Fig. 1**). In comparison to state-of-the-art methods, we evaluate the calibration of PIP through simulations with various genetic architecture settings. We then compare the performance of identifying causal variants, with respect to mapping precision, credible set size, replication rate of discovery in an independent sample, and out-of-sample prediction using fine-mapped variants. Moreover, we develop a LD-based method to construct local credible sets (LCSs), where a α-LCS represents a minimal set of SNPs in high LD that capture a causal variant with a posterior probability of α, and estimate the proportion of SNP-based heritability explained by LCSs. To characterise the overall fine-mapping power in the current study, we propose a concept of global credible sets (GCSs), with a α-GCS representing a minimal set of genome-wide SNPs that capture α% of all causal variants for the trait. Furthermore, leveraging the genetic architecture estimated from SBayesRC, we develop a method to predict the power of fine-mapping and variance explained by the identified variants in prospective studies. With this method, we can estimate the minimal sample size required for identifying a desired proportion of causal variants or those variants explaining a desired proportion of the SNP-based heritability (h^2_{SNP}). Finally, we apply

SBayesRC to the UK Biobank (UKB) data with 13 million SNPs to identify putative causal variants for 599 complex traits and diseases and compare the fine-mapping results using 48 well-powered traits from 6 categories.

Results

Method overview

We selected SBayesRC as the GBMM for GWFM (**Fig. 1**), as it has been shown to outperform other GBMMs in polygenic prediction[27]. SBayesRC is a hierarchical multi-component mixture model, where LD between SNPs is better modelled by matrix factorisation and functional genomic annotations are fitted jointly with the summary statistics in a unified computational framework (**Methods**). To optimize its performance for fine-mapping, we implemented an algorithm to automatically determine the number of mixture components in the model (**Methods**). In contrast to the existing fine-mapping methods, we fit all SNPs simultaneously and employed MCMC sampling to obtain the joint posterior distribution of model parameters and PIPs (**Supplementary Table 1**). In each MCMC iteration, we sampled a dummy variable for each SNP to indicate whether the SNP had a nonzero effect, conditional on the effects of other SNPs. After MCMC sampling, PIP was calculated as the frequency with which the SNP had nonzero effects across the iterations (**Methods**).

A high PIP value provides evidence of a causal variant. However, a causal variant may not have a high PIP value if it is in strong LD with other SNPs. For example, if the causal variant is in perfect LD with another SNP, then the PIP is expected to be 0.5 for each variant, regardless of the sample size. Therefore, the CS concept has been introduced to capture causal variants in strong LD with non-causal SNPs[9,33]. It is common to consider CS for SNPs that are close in physical distance, such as within a 100kb window[34,35]. However, we reason that this approach will miss causal variants with SNPs in long-range LD, and therefore proposed a new method to construct LCS based on LD between SNPs (**Methods**). Starting from the SNP with the largest PIP, we aimed to construct a α-LCS for each "free" SNP (SNP that has not been included in any LCSs), by first selecting other "free" SNPs in high LD ($r^2 > 0.5$) and then summing over their PIPs in a decreasing order until the sum is at least α (a common strategy used in the literature[11,12]). To avoid having too many SNPs with small PIPs in the LCS, we calculated the posterior h^2_{SNP} enrichment probability (PEP), where PEP is the probability that the focal LCS explains more h^2_{SNP} than a random set of SNPs with the same size. The α-LCS was eventually reported if its PEP was greater than 0.7.

In addition to LCS, we proposed another type of CS, GCS. Given the estimated number of causal variants from GBMM (m_c), a α-GCS was computed as the cumulative sum of decreasingly ranked

PIPs that is greater than $\alpha \times m_c$. It can be shown that a α-GCS is expected to cover α% of all causal variants for the trait (**Methods**), with the size of α-GCS reflecting the power of identifying the causal variants given the data (the higher the power, the smaller the α-GCS size). Moreover, from the MCMC samples of SNP effects, we estimated the proportion of h^2_{SNP} explained by the LCSs and GCS (**Methods**).

Based on m_c and the distribution of causal effect sizes estimated from GBMM, we developed a method to predict the power and the proportion of h^2_{SNP} explained by the fine-mapped variants, given a sample size (**Methods** and **Supplementary Note**). This method allows us to estimate the minimal sample size required to achieve a desired power of identifying all causal variants or identifying the causal variants that explain a desired proportion of h^2_{SNP} of the trait, using the ancestry-specific fine-mapping result. Our method to predict fine-mapping power is analytically tractable and has been implemented in a publicly available online tool (https://sbayes.pctgplots.cloud.edu.au/shiny/power/).

We compared SBayesRC to several state-of-the-art fine-mapping methods, including FINEMAP[12], SuSiE[11], FINEMAP-inf[13], SuSiE-inf[7], and PolyFun+SuSiE[19], as well as another GBMM, SBayesC (i.e., two-component SBayesR[24]). All these methods assume a point-normal mixture prior for the SNP effects (**Methods** and **Supplementary Table 1**). A full list of acronyms used in this study can be found in **Supplementary Table 2.**

Calibration of fine-mapping methods under various genetic architectures

We performed extensive genome-wide simulations to calibrate different fine-mapping methods under various genetic architectures, using 100,000 individuals with ~1 million HapMap3 SNPs from the UKB[36]. We started by simulating a sparse genetic architecture, where 1% SNPs were randomly chosen as causal variants, with their effects sampled from a normal distribution, contributing 50% of the phenotypic variance. In this simulation, the data-generative model is consistent with the model used in SBayesC as well as the local fine-mapping methods in comparison. To challenge these methods, we simulated additional two complex genetic architectures (**Methods**). One was a large-effects architecture, where 10 random causal variants contributed 10% of the phenotypic variance and the remaining causal variants contributed 40%. Another complex genetic architecture was based on the sparse architecture but allowed for extensive LD between causal variants and SNP markers. This was achieved by sampling the causal variants only from SNPs in the high LD and high minor allele frequency (MAF) group, and therefore referred to as LD-and-MAF-stratified (LDMS) architecture. To calibrate each fine-

mapping method, we evaluated how well the reported PIPs are consistent with the actual fraction of causal variants, i.e., the true discovery rate (TDR).

Results showed that overall, the GWFM methods had the best calibration, the enhanced region-specific methods with an infinitesimal effect (i.e., FINEMAP-inf and SuSiE-inf) the second, and the standard region-specific methods (i.e., FINEMAP and SuSiE) the worst (**Fig. 2**). Under the sparse genetic architecture, PIPs from SBayesRC/SBayesC were in strong concordance with the TDR across its full spectrum (**Fig. 2a**). The concordance was reasonably good for SuSiE-inf and FINEMAP-inf, although with a trend of deflation in SNPs with low PIP, whereas for SuSiE and FINEMAP, even in SNPs with high PIPs, a notable inflation was observed, indicating a lack of control of FDR (=1-TDR) (**Fig. 2b-c**). When the large-effects or LDMS architecture was used, the assumption in the point-normal BMM was violated in a way that the causal effects did not come from a single normal distribution or that the causal variants were not randomly distributed across the genome. Consequently, PIPs from the point-normal BMM were no longer accurately tracking the true probabilities of causality. When the LDMS architecture was used, the FDR was even more poorly controlled in these methods (**Fig. 2d-i**). However, when SBayesRC was used, with LD and MAF bins as annotations, the strong concordance between PIP and TDR held in various architectures, although none of these architectures matched exactly with the analytic model assumed in SBayesRC.

In conclusion, the region-specific fine-mapping methods tended to have inflated FDR when the model assumptions were not met. In contrast, SBayesRC produced robust PIPs that were well calibrated under various genetic architectures.

Local and global credible sets

In addition to individual SNP PIP, CS is another critical statistic in fine-mapping. Here, we assessed the performance of SBayesRC in identifying LCS and GCS. For each LD block, we computed an α-LCS that contains at least a causal variant with a probability of α and is enriched in h^2_{SNP} (PEP > 0.7). We first evaluated the true discovery rate for identifying LCS, defined as the actual fraction of the LCS with at least a causal variant. The simulation result showed that the SBayesRC has a similar TDR to SuSiE-inf, which had the best PIP calibration among the region-specific fine-mapping methods (**Fig. S1a-c**). However, SBayesRC was significantly more powerful (**Fig. 3a-c**) and had a remarkably smaller LCS size than SuSiE-inf at the same α threshold (**Fig.**

3d-f). For instance, when α=0.9, SBayesRC outperformed SuSiE-inf by up to 221% improvement in power and 41% reduction in LCS size across the three genetic architectures.

The α-GCS is expected to cover α proportion of the causal variants across the genome (**Methods**). Under various simulation scenarios, our GCS accurately represented the true proportion of causal variants (**Fig. S2a-c**), in contrast to the significant deflation observed with SBayesC (**Fig. S2a-c**). Furthermore, we observe a good agreement between estimated and observed power at any given PIP threshold from SBayesRC (**Fig. S2d-f**). Additionally, SBayesRC gave an unbiased estimate for the proportion of h^2_{SNP} explained by the GCS SNPs, regardless of the given α value, under various scenarios (**Fig. S2g-i**).

Improved mapping precision for identifying causal variants

Our simulation results have shown that SBayesRC had the best calibration even under the architecture that matched with the assumed model for the region-specific fine-mapping methods. We next quantified the mapping precision of these methods. The mapping precision was defined as the distance of the identified variant that passed a given PIP threshold to the nearest causal variant. Hence, the smaller the distance, the higher the mapping precision, e.g., the distance is zero if the causal variant itself is identified. Results from the sparse architecture simulation showed that 97.8% of SBayesRC identified SNPs with PIP > 0.9 were the causal variants, and 99% significant SNPs were located within 16.4kb distance to the causal variants (**Fig. 4a**). With the same PIP threshold, 95.5% and 94.3% of SuSiE-inf and FINEMAP-inf identified SNPs were the causal variants, slightly higher than that of 95.3% and 94.0% from SuSiE and FINEMAP, with 99% significant SNPs located within 25.8kb (SuSiE-inf) and 31.3kb (FINEMAP-inf) to the causal variants, compared to that of 32.7kb and 36.7kb for SuSiE and FINEMAP, respectively. In conclusion, given a PIP threshold of 0.9, SBayesRC led to an at least 2% increase in TDR and a 64% (16.4kb/25.8kb) reduction in the distance to the causal variants, both indicating improved mapping precision compared to the existing methods. We also ran a LD block-wise SBayesC analysis, with model parameters estimated from each region separately. Our result showed that the mapping precision remained notable higher than the competing region-specific fine-mapping methods (**Fig. S3**).

In the simulation with large-effects architecture, the mapping precision for all methods decreased due to the decrease of average per-SNP heritability (from 0.5/10,000 to 0.4/9,990). However, SBayesRC still had the highest precision among all methods (**Fig. 4b**). In the simulation with LDMS architecture, SBayesRC demonstrated a substantially higher mapping precision than the other methods (**Fig. 4c**), likely because SBayesRC allowed the model to weigh SNPs differentially

244 based on their LD and MAF property so that the causal variants were better identified.
245 Furthermore, we compared SBayesRC to Polyfun+SuSiE[19], which is a stepwise method that
246 accounts for the effect size stratification (by LD and MAF annotations) through a prior estimated
247 from stratified LD score regression (S-LDSC)[37,38]. Indeed, Polyfun+SuSiE improved the mapping
248 precision compared to the region-specific methods, but was still significantly inferior to
249 SBayesRC (**Fig. 4c**). These simulation results suggested that SBayesRC is a reliable method for
250 GWFM and can substantially improve the mapping precision of identifying causal variants.
251
252 **Improved replication rate of identification with less bias in estimation**
253 In real data analysis, direct evaluation of mapping precision is not feasible, because which
254 variants have causal effects on a trait are often unknown. Alternatively, we can evaluate the
255 replication rate of the identified variants using an independent sample[13]. Here, we define the
256 replication rate as the proportion of variants with a significant PIP (e.g., PIP > 0.9) from the GWAS
257 sample to be repeatedly identified in an independent (replication) sample with the same or a
258 smaller PIP threshold. It is expected that the method that identifies most causal variants from the
259 GWAS sample will have the highest replication rate, as the false positives are unlikely to be
260 replicated.
261
262 We performed simulations using the UKB samples of European ancestry and split samples into
263 independent datasets for discovery and replication. Putative causal variants were identified at
264 the PIP threshold of 0.9 in the GWAS data (n=100,000). We then quantified the replication rate of
265 the putative causal variants at different significance thresholds in two replication datasets
266 (n=100,000 and 200,000). Using SBayesRC, roughly 33% of identified SNPs can be replicated at
267 PIP > 0.9 when replication n = 100,000, and the replication rate increased to 71% when the
268 replication sample size was doubled (**Fig. 4d**). It may seem counter-intuitive that only a fraction
269 of SNPs was replicated despite using the same PIP threshold of 0.9 in both the discovery and
270 replication datasets. This discrepancy is because there exists a sampling variation in the causal
271 variants identified from distinct samples. As expected, the replication rate increased when using
272 a lower threshold for replication, e.g., with PIP > 0.1, 79.6% of the identified SNPs can be
273 replicated when replication n=100,000. Compared to other methods, SBayesRC demonstrated
274 significantly higher replication rate at each of the PIP thresholds, while differences among the
275 other four methods were small. We also quantified the replication rate in the reverse case where
276 the GWAS sample size was 200,000 but the replication sample size was only 100,000, to mimic
277 the reality that the sample size of replication data is often much smaller than that of discovery. In
278 this scenario, we found that 19% of the identified SNPs can be replicated at PIP > 0.9 using

SBayesRC, and the replication rate of SBayesRC remains significantly higher than that of other methods at each PIP threshold (**Fig. S4a**).

We then assessed the replication rate in the UKB height by constructing different discovery and replication datasets as in the simulation. The results were consistent with the observations from the simulation study (**Fig. 4e** and **Fig. S4b**). Compared to the region-specific methods, SBayesRC improved the replication rate by 11.3% (compared to FINEMAP) and by 1.2% (compared to SuSiE-inf) at PIP > 0.9 when replication *n*=100,000, and improved the replication rate by 19% (compared to FINEMAP) and by 3.5% (compared to SuSiE-inf) when replication sample size was doubled.

Moreover, we examined the bias in effect size estimates of putative causal variants identified from fine-mapping, through regressing their marginal effect sizes estimated from the replication samples on the joint effect sizes estimated from the GWAS sample (the regression slope is expected to be one for an unbiased estimation). In the simulation and UKB height analyses, the regression slope from SBayesRC was 0.978 and 0.974, respectively, superior to all the other methods (**Fig. 4f-g**), likely due to the genetic architecture was estimated simultaneously in SBayesRC but was preset or estimated locally in other fine-mapping methods.

These analyses suggested that the identified SNPs from SBayesRC are more likely to be causal because of the relatively high replication rate in independent samples and the negligible bias in effect size estimation, compared to the other methods.

Improved prediction accuracy using fine-mapped variants

Another approach to evaluate the results of fine-mapping is to conduct an out-of-sample prediction using the fine-mapped variants. Since all the Bayesian methods used in this study provide the posterior mean of SNP effects, we computed polygenic scores (PGS) based on the identified variants from each of the methods and evaluated the prediction accuracy in a validation sample. We split the 100K samples into 95K training samples to perform the fine-mapping analysis using all these Bayesian methods and predicted the phenotype for 5K independent individuals as validation samples. We found that overall, SBayesRC had a higher prediction accuracy compared to the other methods, outperforming them by at least ~17% at a PIP threshold of 0.9, with a relatively smaller number of SNPs included in the predictor (**Fig. 5a**). This is consistent with the result that SBayesRC resulted in a lower FDR than the other methods (**Fig. 2a**).

We further compared the performance of these methods by trans-ancestry prediction in real traits. Specifically, we used the fine-mapped variants and estimated posterior effect sizes obtained from the UKB individuals of the European (EUR) ancestry to predict the phenotypes in three other ancestries in UKB: African (AFR), East Asian (EAS) and South Asian (SAS). We selected 6 complex traits that had at least 50 identified SNPs at a PIP threshold of 0.9. We compared the performance of trans-ancestry prediction between SBayesRC and SuSiE-inf, because SuSiE-inf has exhibited a superior performance compared to the others. The result showed that compared to SuSiE-inf, SBayesRC improved the trans-ancestry prediction accuracy using fine-mapped variants, with a nearly 10-fold increase in the mean relative prediction R^2 ($\frac{R^2_{\mathrm{SBayesRC}}-R^2_{\mathrm{SuSiE-inf}}}{R^2_{\mathrm{SuSiE-inf}}}$) across traits and ancestries (**Fig. 5b**). We further compared the performance of SBayesRC and SuSiE-inf for trans-ancestry prediction using the identified credible set SNPs. Similar to the comparison result based on fine-mapped variants, SBayesRC improved the trans-ancestry prediction accuracy based on the SNPs in the 90-LCS (**Fig. S5**; 1.7-fold increase on average). Since it is parsimonious to assume that the common causal variants and their effect sizes are mostly shared between ancestries[39,40], we expect to observe a strong concordance in prediction accuracy between EUR and other ancestries using putative causal variants identified from the EUR sample with high confidence. To investigate this, we quantified the transferability of fine-mapped SNPs by computing the ratio of per-SNP prediction accuracy in a hold-out EUR sample to that in a different ancestry. The result showed that on average this relative prediction accuracy of SNP increases with its PIP calculated in the EUR GWAS sample (**Fig. 5c**). These results suggested that SBayesRC has higher power of fine-mapping and higher accuracy of variant effect estimation.

Prediction of fine-mapping power and variance explained

As a unique feature of the GWFM approach, the genetic architecture estimated from SBayesRC provides information to predict the proportion of causal variants identified from fine-mapping (power) and the proportion of h^2_{SNP} explained by these variants (PHE) for future studies (**Methods** and **Supplementary Note**). To evaluate our approach, we computed the predicted values of power and PHE at a spectrum of GWAS sample sizes and projected the outcome of fine-mapping using SBayesRC onto the prediction using data from the simulated trait with sparse architecture, height[36], high density lipoprotein (HDL), schizophrenia (SCZ)[41], and Crohn's Disease (CD)[42]. These traits were selected to represent different genetic architectures (**Fig. 6a-c**). Despite some variations across traits, the outcomes from the fine-mapping analyses were overall consistent with the theoretical predictions (**Fig. 6d,e**). While our theoretical prediction does not

348 model LD between SNPs, the extent to which the observed values were consistent with the
349 predicted suggests that LD had been effectively, albeit not perfectly, accounted for by our LCSs.
350
351 Take SCZ for example. Using the latest GWAS summary statistics from the psychiatric genomics
352 consortium dataset[41], we identified 13 SNPs and 222 credible sets, collectively explaining 3.9%
353 of h^2_{SNP} at the liability scale[43]. These estimates are highly consistent with our theoretical
354 prediction given the 53,386 cases and 77,258 controls in their study[41], which is equivalent to a
355 sample size of 228,810 on the liability scale (ref[44]; **Methods**). For a prospective study using
356 SBayesRC, we predict that ~180k cases would be required to fine-map 1,000 common causal
357 variants (estimated to be 1.2% of all common causal variants), assuming an equal number of
358 controls and a population prevalence of 0.01 (**Methods**), collectively accounting for about 20%
359 of h^2_{SNP} (**Fig. 6**). With ~550k cases under the same assumption, we will be able to identify 10%
360 causal variants explaining 50% of h^2_{SNP} in SCZ. To fine-map variants accounting for 80% of h^2_{SNP},
361 it was estimated to require 1.4 million cases.
362
363 **Genome-wide fine-mapping in complex traits from UK Biobank**
364 We applied GWFM with SBayesRC to 599 complex traits (597 from the UKB) and developed an
365 online resource to query these fine-mapping results (see Data Availability; **Supplementary**
366 **Table 3**). We selected these 598 UKB traits based on z-score > 4 and high confidence for
367 heritability estimates using LD score regression (https://zenodo.org/records/7186871). To
368 better capture the causal variants, we used 13 million imputed SNPs with functional genomic
369 annotations from Finucane et al.[37]. Here, we focus on discussing the results for 48 complex traits
370 that had sufficient power, including SCZ, CD[41,42] and 46 UKB traits measured in the European
371 ancestry inferred individuals (**Methods**). At the PIP significance threshold of 0.9, we identified
372 2,868 SNPs associated with 48 complex traits, 1,641 of which were not identified by LD clumping,
373 and 22,803 0.9-LCSs with an average size of 8.7 SNPs. On average across these 48 traits, we
374 estimated that although these fine-mapped variants and LCSs only captured 0.75% of the causal
375 variants, they accounted for 17.4% of the h^2_{SNP}.
376
377 Given the estimated genetic architecture for these 48 traits, we applied our theoretical prediction
378 approach to predict the power of prospective GWAS studies. With a GWAS sample size of 2 million
379 individuals, we predict that the average power is 9.5% (**Fig. 7a**) and average PHE is 54.1% (**Fig.**
380 **7b**). The predicted values varied substantially between trait groups. Blood cell traits had both the
381 highest power (29.7%) and highest expected proportion of h^2_{SNP} explained (86.9%), while
382 cognitive traits have the smallest power (16.9%) and smallest expected proportion of genetic
383 variance explained (63.1%). To achieve a PHE of 50%, blood cell traits require a GWAS sample

size of only 1 million individuals, while cognitive traits necessitate a sample size of 5 million individuals, due to the differences in genetic architecture across complex traits. The required sample size increased to 3 million and 10 million for blood cell counts and cognitive traits, respectively, to achieve a PHE of 80% (**Fig. 7b**).

The global credible set α-GCS varied in the credible set size and PHE estimate across traits (**Fig. 7c**). On average, the 0.1-GCS, i.e., covering 10% of causal variants, consisted of 1% of the genome-wide SNPs, which explained 31.8% of the h^2_{SNP} (**Fig. 7d**). Among the analysed complex traits, diseases had the largest GCS sets, requiring 1.9% genome-wide common SNPs to cover 10% common causal variants. In contrast, blood cell traits had the smallest GCS, requiring merely 0.21% genome-wide common SNPs to cover 10% common causal variants. Interestingly, the 0.1-GCS for blood cell traits explained 44.3% of the total genetic variance, compared to 20.8% explained by the GCS for cognitive traits, highlighting the less polygenic genetic architecture of blood cell traits.

Over the 48 complex traits, the number of fine-mapped variants from SBayesRC was highly correlated with the number of identified GWAS loci (**Fig. S6**), ranging from 1 to 489, with an average of 86.2 across complex traits (**Fig. S7**). Compared to the genome-wide SNPs and GWAS identified SNPs after LD clumping, the 2,868 putative causal variants had a significant overrepresentation in functional genomic regions included in the functional genomic annotations[37], such as coding, promoter, and enhancer regions, and were significantly depleted in repressed regions (**Fig. 8a**), suggesting the importance of functional annotations. Of these variants, 651 (22.8%) were in association with more than one complex trait, highlighting the prevalence of pleiotropy in human genome. Moreover, the number of traits that the variant had pleiotropic effects decreased with its minor allele frequency (**Fig. S8**), consistent with that highly pleiotropic variants would be expected to be removed from the population or kept at low frequencies due to natural selection[45].

Functional annotations helped pinpoint the putative causal variants

One notable example is a variant (rs1421085) at the FTO locus, which was identified to be a putative causal variant using SBayesRC for body mass index (BMI), body fat percentage (BFP), hip circumference (HC) and waist circumference (WC). It has been previously validated that this variant plays a causal role in adipocyte thermogenesis regulation[46]. Unlike the results from the standard GWAS where many SNPs in the FTO locus exhibited a signal at genome-wide significance level, our analysis showed that only the known causal variant (rs1421085) was significantly (PIP > 0.9) associated with BMI (**Fig. 8b**). In contrast, applying other methods (SBayesR and SuSiE-inf) without functional annotations to the locus identified the GWAS lead SNP instead, underscoring

420 the importance of incorporating functional annotations. In particular, the annotations of
421 conservation across species, especially primates, helped distinguish the causal variant from the
422 GWAS lead SNP (**Fig. 8c**). Moreover, a secondary signal rs76488452 (PIP=0.85) was identified by
423 SBayesRC, which has not been previously reported but was included in a local credible set of 5
424 SNPs in both SBayesR and SuSiE-inf (**Fig. S9**). We found this SNP resided in a conserved region in
425 primates and was also significant in the COJO analysis (p-value = 1.8×10^{-17}) conditional on the
426 known causal variant. Notably, the secondary signal (rs76488452) was only nominally
427 significantly in the GWAS marginal analysis (p-value = 3.6×10^{-4}), whose trait-increasing allele was
428 in negative LD (r=-0.16) with that of the known causal variant, indicating that this SNP is likely to
429 have a masked effect[47] (estimated masked effect size $b_2 - r * b_1 = 0.02$, consistent with the reported
430 GWAS marginal effect size).
431
432 Another example is from the fine-mapping results for SCZ. We identified 13 SNPs at PIP > 0.9 for
433 SCZ from the latest meta-analysis, 5 of which were the same SNPs that were identified using
434 FINEMAP in their study[41], and all the 8 SNPs identified by FINEMAP were included in our 0.9-
435 LCSs. We recapitulated a missense variant (rs13107325) in *SLC39A8*, a gene highlighted in the
436 latest SCZ analysis for its function in regulating dendritic spine density[48,49]. Furthermore, we
437 identified a secondary variant at the same locus, located in important functional regions (**Fig. 8d-**
438 **e**). Among the 5 novel fine-mapped SNPs that were not identified by FINEMAP with individual
439 PIP, 3 were missense variants (**Fig. S10a-c**). We highlight rs11845184, which is located within
440 *SECISBP2L* (**Fig. S10a**). *SECISBP2L* is highly expressed in brain-related tissues (**Fig. S11**),
441 specifically in differentiating oligodendrocytes, where the SECISBP2L-DIO2-T3 pathway
442 mediates the autonomous regulation of oligodendrocyte differentiation during myelin
443 development[50]. Moreover, we identified novel putative causal variants for CD (**Supplementary**
444 **Table 4**). Using SBayesRC, we fine-mapped 31 variants, of which 10 were missense variants, and
445 all 3 variants identified in the previous study using the same data were recapitulated[42]. In
446 addition, compared to a recent exome-wide association study for CD[51], we identified 4 novel
447 genes (*LACC1*, *SLAMF8*, *MAN2B2* and *GPR35*) with missense putative causal variants (**Fig. S12**).
448 These results demonstrated the power of SBayesRC for identifying the plausible causal variants
449 and provide a valuable resource for downstream analysis and functional validation.
450
451 **Discussion**
452 In this study, we comprehensively evaluated the performance of GWFM using SBayesRC by
453 extensive simulation and real data analyses, compared with the existing fine-mapping methods
454 that consider one genomic region at a time. Our results showed that both PIP and CS from
455 SBayesRC were correctly calibrated under various genetic architectures, indicating well

controlled FDR. In contrast, the other methods produced mis-calibrated PIP and CS with inflated FDR, when the genetic architecture did not match the model assumption. While all fine-mapping methods gave a higher mapping precision than that from GWAS[7], SBayesRC had the highest precision across genetic architecture scenarios. Furthermore, in both simulation and real trait analyses, SBayesRC showed significantly higher replication rate and prediction accuracy but less estimation bias in an independent sample using fine-mapped SNPs, compared with the other methods. In the real data analysis, we showed examples where SBayesRC pinpointed the putative causal variants that were missed by the other methods. All of these results indicate that SBayesRC, as a method for GWFM analysis, remarkably improves the identification of causal variants.

We proposed a new LD-based method to compute LCSs and estimate their contribution to the SNP-based heritability. This method overcomes the limitation of existing window-based approaches that causal variants with SNPs in long-range LD would not be captured. In addition to LCSs, SBayesRC allows us to compute a GCS for the trait, which informed the power of identifying the causal variants and the h^2_{SNP} explained by the identified SNPs given the data. This computation requires an unbiased estimation on the total number of the causal variants, which can only be done when analysing all SNPs jointly in the model. The analysis of 48 complex traits showed that although as many as 22,803 variants or LCS were identified, they only captured 0.75% of all common causal variants and contributed 17.4% genetic variance, suggesting many causal variants with very small effects are yet to be discovered (**Fig. 7c,d**).

We have provided a theoretical prediction of fine-mapping power given a sample size and the estimated genetic architecture (**Supplementary Note**). This is useful to inform the experimental design of future fine-mapping studies regarding the sample size required to identify a certain number of causal variants or those explaining a certain proportion of h^2_{SNP}. The robustness of our prediction approach is supported by projecting the outcomes of real data analyses to the landscape of predicted values. For height, based on the UKB data (n = 350k), we predicted that when the sample size increases to 5 million, the number of fine-mapping discoveries would be ~10,000 considering significant PIPs only or ~30,000 considering both significant PIPs and LCSs, explaining up to 95% of the genetic variance (**Fig. 6**). This prediction is consistent with the finding of a recent GWAS with 5 million individuals, which reported 12,111 independently significant SNPs identified from COJO accounting for nearly all of the common SNP-based heritability in height[40].

While the concept of credible set has evolved over time[11,12,52], it is still common to focus on individual SNP PIP in the downstream analysis, probably because the CS include too many SNPs

492 to follow up. Our study provided important implications regarding this issue. First, CS may miss
493 the true causal variant if not all possible causal variants are fitted in the model, underscoring the
494 importance of considering all common SNPs in the fine-mapping analysis. Second, our GWFM
495 approach can reduce the credible size, as shown in both simulation and real trait analysis (only
496 ~8.7 SNPs per credible set), facilitating the consideration of CS in practice. Third, in the presence
497 of complete LD between SNPs and the causal variants, the PIP of a causal variant may never be
498 significant regardless of sample size, but leveraging functional genomic annotations may help
499 distinguish causal from non-causal variants. In this sense, genomic annotations play a greater role
500 than the increase of GWAS sample size.
501
502 The advantages of SBayesRC over the region-specific fine-mapping methods arise from the
503 following aspects. First, SBayesRC involves a genome-wide analysis fitting all SNPs jointly.
504 Compared to the region-specific analysis, genome-wide analysis accounts for long-range LD and
505 utilises all SNPs to estimate the genetic architecture, thereby improving fine-mapping. Of note,
506 even when the same LD blocks are used in both types of analyses, the latter is still superior
507 because of the better estimation of genetic architecture parameters from genome-wide SNPs.
508 Second, SBayesRC assumes a more realistic distribution of SNP effects through using MAF/LD
509 groups along with other functional annotations. In addition, the impact of annotations on the SNP
510 effect distribution is estimated within the same model, fostering a formal Bayesian learning
511 process. The existing fine-mapping methods, however, either do not leverage annotation data or
512 lack a unified framework for the joint analysis with GWAS data. Third, SBayesRC utilises MCMC
513 sampling to estimate model parameters and PIPs, which is asymptotically exact. Both FINEMAP
514 and FINEMAP-inf use shotgun stochastic search, while both SuSiE and SuSiE-inf use variational
515 Bayes to compute the Bayes factors for the causal models and therefore the PIPs. It has been
516 previously shown that MCMC sampling generally leads to a higher accuracy of capturing the
517 posterior distribution than the other approximation approaches[30]. To further justify our choice
518 of SBayesRC as the method for GWFM, we ran SBayesRC within each block separately and
519 quantified the mapping precision. We found that the mapping precision decreased compared to
520 the genome-wide SBayesRC but remained higher than the other methods (**Fig. S2**). For example,
521 99% of SNPs identified by the region-specific SBayesRC were located within 23.1Kb to causal
522 variants, compared to the number of 16.4Kb for the genome-wide SBayesRC and 25.8Kb for
523 SuSiE-inf.
524
525 We note several limitations of this work. First, there are certainly more complicated scenarios
526 about effect size distribution for causal variants that have not been investigated in our
527 simulations. However, to our knowledge, SBayesRC is one of the most flexible models to

accommodate various scenarios because it assumes a multi-component Gaussian mixture, and we have further allowed the method to automatically choose the number of mixture components. Second, unlike an individual-level model where each PIP is calculated conditional on the effects from all other SNPs jointly, SBayesRC is a summary-level model where LD between LD blocks is ignored so that SNPs beyond the region contribute to the PIP only through the mixture distribution of SNP effects. Third, we only applied our method to the GWAS summary data from relatively homogenous populations (inferred European ancestry) and the robustness of the methods on GWAS data based on trans-ancestry meta-analyses is not investigated. Fourth, SBayesRC requires the LD information estimated from a cohort that matches with GWAS ancestry without substantial sampling variation. Fifth, to create the credible set, a threshold of 0.5 was arbitrarily chosen to define a set of SNPs in high LD. Latest methodological advancements in Bayesian hypothesis tests based on hierarchical clustering can be used to relax this condition[53]. Sixth, the prediction of mapping power is based on the genetic architecture estimates given a SNP set. However, the SNP set may change with the sample size (e.g., more common SNPs to be observed in a larger sample size), which may affect the polygenicity and SNP-based heritability. Despite these limitations, our study provides a robust and versatile GWFM framework for identifying causal variants, highlighting the advantages of this approach over existing region-specific fine-mapping methods. With its capacity to enhance mapping power in the current study and to predict mapping power for future studies, we believe GWFM using a state-of-the-art GBMM will become the preferred method for analysing complex traits.

Methods

Ethics approval

The University of Queensland Human Research Ethics Committee B (2011001173) provides approval for analysis of human genetic data used in this study on the high-performance cluster of the University of Queensland.

Low-rank GBMM

We used state-of-the-art GBMM that employed a low-rank model to improve computational efficiency and robustness[27]. As described below, the low-rank GBMM can be derived from the individual-level model. Consider a multiple linear regression of phenotypes on genotypes:

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \tag{1}$$

where $\mathbf{y}$ is an $n \times 1$ vector of complex trait phenotypes and $\mathbf{X}$ is an $n \times m$ matrix of mean-centred genotypes, $\boldsymbol{\beta}$ is $m \times 1$ vector of SNP effects on the trait, and $\mathbf{e}$ is $n \times 1$ vector of residuals with $\mathrm{var}(\mathbf{e}) = \mathbf{I}\sigma_e^2$. Let

$$\mathbf{K} = \boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{U}'\mathbf{X}'n^{-1} \tag{2}$$

where $\boldsymbol{\Lambda}$ and $\mathbf{U}$ are diagonal matrix of eigenvalues and matrix of eigenvectors for the LD correlation matrix $\mathbf{R} = \mathbf{X}'\mathbf{X}/n$, respectively. It follows that $\mathbf{K}'\mathbf{K} = \mathbf{P}n^{-1}$, where $\mathbf{P}$ is the projection matrix of $\boldsymbol{y}$ on the column space of $\mathbf{X}$, and $\mathbf{K}\mathbf{K}' = \mathbf{I}n^{-1}$. Multiplying both sides of Eq (2) by $\mathbf{K}$ gives

$$\mathbf{K}\mathbf{y} = \mathbf{K}\mathbf{X}\boldsymbol{\beta} + \mathbf{K}\boldsymbol{e} \tag{3}$$

or

$$\mathbf{w} = \mathbf{Q}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{4}$$

When only the top q principal components of $\mathbf{R}$ are used, the dimension of $\mathbf{w}$ is $q \times 1$ and $\mathbf{Q}$ is $q \times m$. Since $q \ll n$, this model would have a substantially lower rank than Eq (1), improving the computational efficiency for the estimation of $\boldsymbol{\beta}$. With a recognition of $\mathbf{b} = \mathbf{X}'\boldsymbol{y}/n$ is the GWAS marginal effect estimates, $\mathbf{w}$ can be directly computed from the GWAS summary statistics. In practice, we compute $\mathbf{w}$ and $\mathbf{Q}$ within quasi-independent LD blocks in the human genome. With this low-rank model, we can estimate $\boldsymbol{\beta}$ for all common variants jointly through considering $\boldsymbol{\beta}$ as random effects. In addition, this model allows a direction estimation of the residual variance, as $\mathrm{var}(\boldsymbol{\epsilon}) = \mathbf{I}\sigma_e^2 n^{-1}$, which can be used as a nuisance parameter to accommodate heterogeneity in the summary statistics and LD reference[27].

SBayesC and SBayesRC

GBMMs are flexible in the specification of the prior distribution of SNP effects. In SBayesC, the prior for the effect size of variant j is,

$$\beta_j \sim N\left(0,\ \sigma_\beta^2\right)\pi + \phi(1-\pi) \tag{5}$$

where σ_β^2 is the common variance across all the causal variants, π is the proportion of SNPs with nonzero effects, and ϕ is a point mass at zero. Both σ_β^2 and π are considered as unknown, with a scaled inverse chi-squared prior distribution and a uniform prior distribution, respectively[27].

SBayesRC[27] is an extension of SBayesR[24], which allows for a more realistic prior distribution of SNP effects by assuming a multiple component mixture distribution

$$\beta_j \sim \sum_{k=1}^{5} \pi_k N\left(0,\ \gamma_k \sigma_g^2\right) \tag{6}$$

where $\boldsymbol{\gamma} = \left(\gamma_1, \gamma_2, \gamma_3, \gamma_{4,} \gamma_5\right)' = (0, 1 \times 10^{-5}, 1 \times 10^{-4}, 1 \times 10^{-3}, 1 \times 10^{-2})'$ are the prespecified coefficients to constrain the variance in each effect size distribution with respect to the total genetic variance (σ_g^2), and π_k is the prior probability for the SNP effect belong to the kth distribution. To further account for the stratification of causal variants and their effects regarding functional annotations, SBayesRC assumes a SNP-specific prior probability of distribution membership, π_{jk}, depending on the annotations at each SNP, through a generalised linear model. Specifically,

$$f\left(\pi_{jk}\right) = \mu_k + \sum_{l=1}^{c} A_{jl}\, \alpha_{kl} \tag{7}$$

where $f(\cdot)$ is the probit link function, μ_k is the intercept, A_{jl} is the value of annotation l on SNP j (either binary or continuous annotations), and α_{kl} is the effect of annotation l on the prior probability of the SNP effect belonging to the kth distribution. Details of the method can be found in ref[27].

Calculation of PIP

We assessed the strength of joint association of each SNP using the posterior inclusion probability (PIP), i.e., the probability of a SNP being included with a nonzero effect in the model, given the data. Let δ_j be the indicator variable for the distribution membership for SNP j, with $\delta_j = 1$ indicating a null effect and $\delta_j = 2, \dots, K$ indicating a nonnull component. We computed PIP for SNP j as

$$\mathrm{PIP}_j = 1 - \Pr\left(\delta_j = 1 \middle| y\right) = 1 - \frac{f(y|\delta_j=1)\pi_1}{\sum_{k=1}^{J} f(y|\delta_j=k)\pi_k} \tag{8}$$

The likelihood function when $\delta_j = 1$ are

$$f(y|\delta_j = 1) \propto \exp\left\{-\frac{y_c' y_c}{2\sigma_e^2}\right\} \tag{9}$$

where y_c is the adjusted y for all other effects except that for SNP j.

The likelihood function when $\delta_j = k$ is

$$f(y|\delta_j = k) \propto \exp\left\{-\frac{y_c' y_c}{2\sigma_e^2}\right\} \lambda_k^{\frac{1}{2}} C_k^{-\frac{1}{2}} \exp\left\{\frac{r^2}{2C_k}\right\} \tag{10}$$

where $\lambda_k = \frac{\sigma_e^2}{\gamma_k \sigma_g^2}$, $C_k = n + \lambda_k$, $r = X'y_c = X'e + n\beta$. A full derivation of equation above can be found in the supplementary Note.

For all GBMM analyses in this study, we ran Markov chain Monte Carlo (MCMC) sampling for 10,000 iterations with the first 2,000 samples as burn-in and we used the posterior mean over 8,000 posterior samples to estimate $\boldsymbol{\pi}$ and PIPs.

Automatic determination of mixture components

The standard SBayesRC requires specification of the number of mixture components before the analysis. It has been shown that the performance of polygenic prediction is robust to the number of mixture components[27]. However, this may be a problem for fine-mapping if a small effect component is unnecessarily included, where null SNPs are fitted by chance to explain negligible variance. This is because these SNPs may affect the distribution of PIPs and cause a bias in the estimation of the number of causal variants. In this study, we have allowed the method to automatically determine the number of mixture components for SBayesRC. The procedure started with running SBayesRC using the default setting of five mixture components. After 500 iterations of MCMC, the smallest component would be removed if the genetic variance explained by the SNPs in this component were less than half of that in the second smallest component. This procedure was repeated until no component was removed from the model. The rationale is that in most complex traits, due to the action of negative selection, most variation is attributed to variants with small effects[31,32,54]. Hence, if the smallest component is capturing true genetic effects, it should contribute to a significant proportion of variance, unlikely to be substantially lower than the second smallest component.

Local and global credible sets

Similar to prior work[14], we defined the local credible set SNPs as the minimum set of SNPs that contains at least one causal variant with a probability of α. To identify the α-LCS, we ranked SNPs based on their PIPs and constructed candidate credible set for each "free" SNP which was not in any LCSs. For the focal SNP, the candidate credible set was created by including "free" SNPs in high LD ($r^2 > 0.5$) with a focal SNP and computed the α-LCS by summing over PIPs in a decreasing order until the sum is at least α. This process was iteratively repeated until all SNPs were exhausted. For each α-LCS, we calculated an LCS posterior SNP-heritability enrichment probability, where PEP is the probability that the focal LCS explains more h^2_{SNP} than a random set of SNPs with the same size. We reported all the 0.9-LCS with PEP > 0.7 for each LD block. The true discovery rate was quantified as the proportion of identified α-LCS containing at least one causal

variant, and the power was calculated as the proportion of simulated causal variants included in the identified α-LCS.

Analogous to the LCS, which identifies a set of SNPs that capture a causal variant with a probability of α, the GCS identifies a set of SNPs that capture all causal variant with a probability of α, which is equivalent to finding a set of SNPs that capture α% of the causal variants. We computed the α-GCS as the cumulative sum of decreasingly ranked PIPs that was greater than $\alpha \times m_c$, where m_c was the estimated number of causal variants from GBMM for the trait. The α-GCS is expected to cover α% of all causal variants for the trait, i.e., the power of identifying the causal variants given the data (**Supplementary Note**).

Estimation of power and variance explained given the data

For the identified SNPs using individual PIP or credible set, we estimated the power of identifying the causal variants given the data at a given threshold α,

$$\mathrm{TPR}_{\alpha} = \frac{\sum_j [\mathrm{PIP}_j | \ \mathrm{PIP}_j \geq \alpha]}{M\pi} \tag{11}$$

A formal derivation is given in the **Supplementary Note**.

Moreover, we estimated the proportion of SNP-based heritability explained (PHE) by LCSs. Specifically, we computed PHE for a focal set (i) of SNPs in each MCMC iteration using the sampled values of SNP effects,

$$\mathrm{PHE}_{mcmc_r,i} = \frac{\sum_j \left[\beta^2_{mcmc_r,j} | \ j \in i\right]}{\sum_{m=1}^{M} \left[\beta^2_{mcmc_r,m}\right]} \tag{12}$$

where $\beta_{mcmc_r,j}$ is the sampled effect for SNP j from MCMC iteration r in the scaled genotype unit. Finally, we computed the posterior mean across MCMC iterations as the estimate for PHE_i,

$$\mathrm{PHE}_i = \frac{\sum_r \mathrm{PHE}_{mcmc_r,i}}{L} \tag{13}$$

where L is the total number of MCMC iterations.

Prediction of power and variance explained for prospective studies

We aim to predict the power of identifying a certain proportion of the causal variants in a prospective fine-mapping analysis, given a sample size (n) and the genetic architecture of the trait, when PIP from a GBMM is used as the test statistic. As shown in the **Supplementary Note**, assuming that variance explained by the causal variant is v, the sampling distribution of PIP from the multi-component mixture model, e.g., SBayesRC, is

$$PIP = 1 - \frac{1}{1 + \sum_{k=2}^{5} A_k \exp\{B_k Z\}} \tag{14}$$

where $A_k = \frac{\pi_k}{\pi_1}\lambda_k^{\frac{1}{2}}C_k^{-\frac{1}{2}}$ and $B_k = \frac{n\sigma_e^2}{2C_k}$ are two constants given the genetic architecture parameters ($\boldsymbol{\pi}$ and h_{SNP}^2), with $\lambda_k = \frac{\sigma_e^2}{\gamma_k\sigma_g^2}$ and $C_k = n + \lambda_k$, and Z is a data-dependent random variable following a non-central Chi-square distribution with the non-centrality parameter $NCP = \frac{nv}{\sigma_e^2}$:

$$Z \sim \chi_1^2(\frac{nv}{\sigma_e^2}) \tag{15}$$

Given the threshold of PIP being α for the hypothesis test, the power to detect this causal variant can be calculated as

$$Power_v = \Pr(PIP > \alpha|v) = \int_\alpha^1 f(PIP|v)\mathrm{d}PIP \tag{16}$$

where $f(PIP|v)$ is Eq (14) above. To compute the power for identifying any causal variant, we integrated out v by

$$Power = \int_\alpha^1 \int_0^\infty f(PIP|v)f(v)\mathrm{d}v\mathrm{d}PIP \tag{17}$$

where

$$f(v) = f_\beta(v^{\frac{1}{2}})v^{-\frac{1}{2}} \tag{18}$$

and $f_\beta(\cdot)$ is the distribution of β estimated from the SBayesRC model.

Therefore, given a sample size, the expected number of causal variants identified from fine-mapping is

$$E[NCV] = m(1-\pi_1) \times Power \tag{19}$$

The expected proportion of genetic variance explained by the fine-mapped variants is

$$E[PHE] = m(1-\pi_1)\sigma_g^{-2}\int_0^\infty Power_v \times vf(v)\mathrm{d}v \tag{20}$$

Since it is computationally challenging to obtain an analytical solution, we opted to estimate these quantities through Monte Carlo simulation (**Supplementary Note**).

Disease sample size at the liability scale

For diseases or binary traits, we converted the GWAS summary statistics from the linear mixed model to the liability scale prior to running GBMM. Based on the method in Yang et al.[44], we estimated the sample size at the liability scale that gives an equivalent power to detect a locus affecting a quantitative trait with the same properties,

$$N_{eq} = \frac{i^2\nu(1-\nu)N_{01}}{(1-K)^2} \tag{21}$$

where K is the disease prevalence, ν is the sample prevalence, $i = h/K$ with h being the height of standard normal curve at the truncation point $t = 1 - K$, and N_{01} is the total number of cases and controls. Given the z-score (z_j) from the original GWAS summary data for SNP j, the marginal effect and its standard error at the liability scale can be estimated as following ref[55]

$$SE_j = \sqrt{\frac{1}{2p_j(1-p_j)(N_{eq}+z_j^2)}} \quad (22)$$

$$b_j = z_j \times SE_j \quad (23)$$

where p_j is the minor allele frequency of the SNP.

The results from GBMM using the converted summary statistics will be directly comparable across traits regardless of the sample prevalence and the type of traits. In our prediction analysis of power, we compared results between diseases and quantitative traits based on the equivalent sample size estimated from Eq (23). Similarly, to estimate the number of cases required, in a case-control study with equivalent number of controls, to achieve a certain power, we rearranged the same equation so that

$$N_{cases} = \frac{1}{2}\frac{(1-K)^2 N_{eq}}{i^2 \nu(1-\nu)} \quad (24)$$

Simulations based on imputed genotype data from the UK Biobank

To evaluate the performance of GBMM, we ran simulations using the imputed genotype data from the UK Biobank after quality controls (QC). In this study, we selected 300,000 unrelated individuals and included ~1.2 million HapMap3 SNPs with MAF > 0.01, Hardy-Weinberg equilibrium test $P > 1\times10^{-6}$, genotyping rate > 0.95, and imputation information score > 0.8 for simulations.

We randomly sampled $m_c = 10{,}000$ casual variants from the genome for 100,000 individuals and simulated complex trait phenotypes based on the following model:

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{e} \quad (25)$$

where $\mathbf{X}$ is the genotype matrix for the causal variants, $\beta_i \sim N(0, h^2/m_c)$ and $e_j \sim N(0, var(\mathbf{X}\boldsymbol{\beta})/(1/h^2 - 1))$ with $h^2 = 0.5$ being the proportion of phenotypic variance explained by all the causal variants. To check the robustness of GBMM, we also ran simulations under various settings. For simulations under LD-MAF stratification model, we partitioned all the genome-wide SNPs into four LD and MAF groups (by their median values) and only sampled the 10,000 causal variants in the high LD and high MAF group. For the major gene model simulation, we separated the sampling of effect size for causal variants from two distributions, i.e., 10 random SNPs with effects from $N(0, 0.2 * h^2/10)$ and the rest SNPs with effects from $N(0, 0.8 * h^2/9990)$.

We ran a standard GWAS using the genotypes with the simulated phenotypes under different settings. We then used the GWAS summary data to perform GBMM (SBayesRC[27] and SBayesC[24]) implemented in GCTB, SuSiE[11], FINEMAP[12], SuSiE-inf[13], FINEMAP-inf[13] and PolyFun-S[19] to detect

749 fine-mapped variants and compute corresponding PIPs and effect sizes. We used imputed
750 genotypes from 10,000 random samples from UK Biobank as the LD reference in this study. We
751 repeated the whole process 100 times and then quantified the true discovery rate, mapping
752 precision and replication rate for each method. The mapping precision was computed as the
753 physical distance between the identified SNPs and nearest causal variants.
754
755 **Real data analysis**
756 We analysed 598 UK Biobank complex traits GWAS summary data from Neale's lab (**Data**
757 **Availability**) and the schizophrenia[41] and Crohn's disease[42] GWAS summary data. We selected
758 these 598 traits with z-score > 4 and high confidence for heritability estimates using LD score
759 regression. We used the annotations from baseline model BaseLineLD v2.2[38] and extract the
760 imputed SNPs with MAF > 0.001 and that are in common with the annotations, leading to
761 13,065,104 imputed SNPs passed the quality control. We used 10,000 random samples from the
762 UK Biobank as the LD reference to run the SBayesRC and other region-specific fine-mapping
763 analysis. We further extracted 48 well-powered traits with relatively large sample size (n > 100,
764 000), high heritability (h^2 > 0.1) and at least a fine-mapped SNP at PIP > 0.9.
765
766 For the polygenic score prediction analysis using fine-mapped variants only, we performed
767 quality control for the imputed genotype data provided by the UKB analysis team[36]. We kept SNPs
768 with MAF > 0.01, Hardy-Weinberg Equilibrium test P > 10^{-10}, imputation info score > 0.6 within
769 each ancestry samples. We removed samples with mismatched sex information, samples
770 withdrawn from participation and cryptic related samples following ref[27]. We separate the final
771 UKB dataset into 4 ancestries, European (EUR, N= 347,800), East Asian (EAS, N=2,252), South
772 Asian (SAS, N=9,436) and African (AFR, N=7,006). The phenotypes with continuous values were
773 filtered within the range of mean +/- 7SD and then rank-based inverse-normal transformed
774 within each ancestry and sex group. The GWAS were performed using PLINK2 software[5] with sex,
775 age and first 10 principal component as covariates. Linear regression was used for continuous
776 traits and logistic regression for binary traits.
777
778 **Supplementary Information**
779 Supplementary data include 12 supplementary figures, 4 supplementary table and a
780 supplementary note.
781
782 **Data Availability**
783 Our SBayesRC-enabled genome-wide fine-mapping results for 599 complex traits are available at
784 link (https://sbayes.pctgplots.cloud.edu.au/data/SBayesRC/share/Finemap/v1.0/). The UK

Biobank data are available through formal application to the UK Biobank (http://www.ukbiobank.ac.uk). The GWAS summary data for 598 complex traits in UK Biobank are from http://www.nealelab.is/uk-biobank/. All the other datasets used in this study are available in the public domain.

Code Availability

Summary-data-based genome-wide Bayesian mixture models are implemented in a public available software GCTB at https://cnsgenomics.com/software/gctb/#Download. Methods to compute LCS and GCS have also been implemented in GCTB (https://cnsgenomics.com/software/gctb/#Genome-wideFine-mappinganalysis). Online tool for predicting fine-mapping power: https://sbayes.pctgplots.cloud.edu.au/shiny/power/.

Acknowledgements

This research was supported by Australian National Health and Medical Research Council (1177268, 1113400), the Australian Research Council (FL180100072, DP220101947, DP230101352) and the National Institute of Mental Health (5R01MH121545-05). This study makes use of data from the UK Biobank (project ID: 12505).

Author Contributions

J.Z. conceived and supervised the study. J.Z. and Y.W. developed the methods. J.Z., P.M.V. and Y.W. designed the experiment. Y.W. conducted all analyses with the assistance or guidance from J.Z., Z.Z., L.T., P.M.V., N.R.W. and M.E.G.. Y.W., and J.Z. wrote the manuscript with the participation of all authors. All the authors approved the final version of the manuscript.

Competing Interests

The authors declare no competing interests.

References

1 Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics* **101**, 5-22, doi:10.1016/j.ajhg.2017.06.005 (2017).

2 Visscher, P. M., Brown, M. A., McCarthy, M. I. & Yang, J. Five years of GWAS discovery. *Am J Hum Genet* **90**, 7-24, doi:10.1016/j.ajhg.2011.11.029 (2012).

3 Abdellaoui, A., Yengo, L., Verweij, K. J. H. & Visscher, P. M. 15 years of GWAS discovery: Realizing the promise. *Am J Hum Genet* **110**, 179-194, doi:10.1016/j.ajhg.2022.12.011 (2023).

4 Uffelmann, E. *et al.* Genome-wide association studies. *Nature Reviews Methods Primers* **1**, 59, doi:10.1038/s43586-021-00056-9 (2021).

5 Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics* **81**, 559-575, doi:https://doi.org/10.1086/519795 (2007).

6 Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature Genetics* **44**, 369-375, doi:10.1038/ng.2213 (2012).

7 Wu, Y., Zheng, Z., Visscher, P. M. & Yang, J. Quantifying the mapping precision of genome-wide association studies using whole-genome sequencing data. *Genome Biology* **18**, 86, doi:10.1186/s13059-017-1216-0 (2017).

8 Farh, K. K.-H. *et al.* Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518**, 337-343, doi:10.1038/nature13835 (2015).

9 Schaid, D. J., Chen, W. & Larson, N. B. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nature Reviews Genetics* **19**, 491-504, doi:10.1038/s41576-018-0016-z (2018).

10 Stephens, M. False discovery rates: a new deal. *Biostatistics* **18**, 275-294, doi:10.1093/biostatistics/kxw041 (2016).

11 Wang, G., Sarkar, A., Carbonetto, P. & Stephens, M. A Simple New Approach to Variable Selection in Regression, with Application to Genetic Fine Mapping. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **82**, 1273-1300, doi:10.1111/rssb.12388 (2020).

12 Benner, C. *et al.* FINEMAP: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics* **32**, 1493-1501, doi:10.1093/bioinformatics/btw018 (2016).

13 Cui, R. *et al.* Improving fine-mapping by modeling infinitesimal effects. *Nature Genetics*, doi:10.1038/s41588-023-01597-3 (2023).

14 Zou, Y., Carbonetto, P., Wang, G. & Stephens, M. Fine-mapping from summary data with the "Sum of Single Effects" model. *PLOS Genetics* **18**, e1010299, doi:10.1371/journal.pgen.1010299 (2022).

15 Chen, W. *et al.* Fine Mapping Causal Variants with an Approximate Bayesian Method Using Marginal Test Statistics. *Genetics* **200**, 719-736, doi:10.1534/genetics.115.176107 (2015).

16 Yang, Z. *et al.* CARMA is a new Bayesian model for fine-mapping in genome-wide association meta-analyses. *Nature Genetics* **55**, 1057-1065, doi:10.1038/s41588-023-01392-0 (2023).

855 17 Kichaev, G. *et al.* Integrating Functional Data to Prioritize Causal Variants in Statistical
856 Fine-Mapping Studies. *PLOS Genetics* **10**, e1004722, doi:10.1371/journal.pgen.1004722
857 (2014).
858 18 Kichaev, G. & Pasaniuc, B. Leveraging Functional-Annotation Data in Trans-ethnic Fine-
859 Mapping Studies. *The American Journal of Human Genetics* **97**, 260-271,
860 doi:https://doi.org/10.1016/j.ajhg.2015.06.007 (2015).
861 19 Weissbrod, O. *et al.* Functionally informed fine-mapping and polygenic localization of
862 complex trait heritability. *Nature Genetics* **52**, 1355-1363, doi:10.1038/s41588-020-
863 00735-5 (2020).
864 20 Zhang, Y., Qi, G., Park, J.-H. & Chatterjee, N. Estimation of complex effect-size distributions
865 using summary-level statistics from genome-wide association studies across 32 complex
866 traits. *Nature Genetics* **50**, 1318-1326, doi:10.1038/s41588-018-0193-x (2018).
867 21 Erbe, M. *et al.* Improving accuracy of genomic predictions within and between dairy cattle
868 breeds with imputed high-density single nucleotide polymorphism panels. *J Dairy Sci* **95**,
869 4114-4129, doi:10.3168/jds.2011-5019 (2012).
870 22 MacLeod, I. M. *et al.* Exploiting biological priors and sequence variants enhances QTL
871 discovery and genomic prediction of complex traits. *BMC Genomics* **17**, 144,
872 doi:10.1186/s12864-016-2443-6 (2016).
873 23 Habier, D., Fernando, R. L., Kizilkaya, K. & Garrick, D. J. Extension of the bayesian alphabet
874 for genomic selection. *BMC Bioinformatics* **12**, 186, doi:10.1186/1471-2105-12-186
875 (2011).
876 24 Lloyd-Jones, L. R. *et al.* Improved polygenic prediction by Bayesian multiple regression on
877 summary statistics. *Nature Communications* **10**, 5086, doi:10.1038/s41467-019-12653-
878 0 (2019).
879 25 Cleveland, M. A., Forni, S., Deeb, N. & Maltecca, C. Genomic breeding value prediction using
880 three Bayesian methods and application to reduced density marker panels. *BMC*
881 *Proceedings* **4**, S6, doi:10.1186/1753-6561-4-S1-S6 (2010).
882 26 Moser, G. *et al.* Simultaneous Discovery, Estimation and Prediction Analysis of Complex
883 Traits Using a Bayesian Mixture Model. *PLOS Genetics* **11**, e1004969,
884 doi:10.1371/journal.pgen.1004969 (2015).
885 27 Zheng, Z. *et al.* Leveraging functional genomic annotations and genome coverage to
886 improve polygenic prediction of complex traits within and between ancestries. *Nature*
887 *Genetics*, doi:10.1038/s41588-024-01704-y (2024).
888 28 Xiang, R. *et al.* Genome-wide fine-mapping identifies pleiotropic and functional variants
889 that predict many traits across global cattle populations. *Nature Communications* **12**, 860,
890 doi:10.1038/s41467-021-21001-0 (2021).

29 Shrestha, M. *et al.* Evaluation of Bayesian Linear Regression Models as a Fine Mapping tool. *bioRxiv*, 2023.2009.2001.555889, doi:10.1101/2023.09.01.555889 (2024).

30 Carbonetto, P. & Stephens, M. Scalable Variational Inference for Bayesian Variable Selection in Regression, and Its Accuracy in Genetic Association Studies. *Bayesian Analysis* **7**, 73-108, 136 (2012).

31 Zeng, J. *et al.* Signatures of negative selection in the genetic architecture of human complex traits. *Nature Genetics* **50**, 746-753, doi:10.1038/s41588-018-0101-4 (2018).

32 Zeng, J. *et al.* Widespread signatures of natural selection across human complex traits and functional genomic categories. *Nature Communications* **12**, 1164, doi:10.1038/s41467-021-21446-3 (2021).

33 Spain, S. L. & Barrett, J. C. Strategies for fine-mapping complex traits. *Human Molecular Genetics* **24**, R111-R119, doi:10.1093/hmg/ddv260 (2015).

34 Gong, Y., Greenbaum, J. & Deng, H. W. A statistical approach to fine-mapping for the identification of potential causal variants related to human intelligence. *J Hum Genet* **64**, 781-787, doi:10.1038/s10038-019-0623-3 (2019).

35 Kamiza, A. B. *et al.* Multi-trait discovery and fine-mapping of lipid loci in 125,000 individuals of African ancestry. *Nature Communications* **14**, 5403, doi:10.1038/s41467-023-41271-0 (2023).

36 Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-209, doi:10.1038/s41586-018-0579-z (2018).

37 Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics* **47**, 1228-1235, doi:10.1038/ng.3404 (2015).

38 Gazal, S. *et al.* Functional architecture of low-frequency variants highlights strength of negative selection across coding and non-coding annotations. *Nature Genetics* **50**, 1600-1607, doi:10.1038/s41588-018-0231-8 (2018).

39 Hou, K. *et al.* Causal effects on complex traits are similar for common variants across segments of different continental ancestries within admixed individuals. *Nature Genetics* **55**, 549-558, doi:10.1038/s41588-023-01338-6 (2023).

40 Yengo, L. *et al.* A saturated map of common genetic variants associated with human height. *Nature* **610**, 704-712, doi:10.1038/s41586-022-05275-y (2022).

41 Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502-508, doi:10.1038/s41586-022-04434-5 (2022).

42 de Lange, K. M. *et al.* Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nature Genetics* **49**, 256-261, doi:10.1038/ng.3760 (2017).

927 43 Lee, Sang H., Wray, Naomi R., Goddard, Michael E. & Visscher, Peter M. Estimating Missing
928 Heritability for Disease from Genome-wide Association Studies. *The American Journal of*
929 *Human Genetics* **88**, 294-305, doi:10.1016/j.ajhg.2011.02.002 (2011).
930 44 Yang, J., Wray, N. R. & Visscher, P. M. Comparing apples and oranges: equating the power
931 of case-control and quantitative trait association studies. *Genet Epidemiol* **34**, 254-257,
932 doi:10.1002/gepi.20456 (2010).
933 45 Novo, I., López-Cortegano, E. & Caballero, A. Highly pleiotropic variants of human traits
934 are enriched in genomic regions with strong background selection. *Human Genetics* **140**,
935 1343-1351, doi:10.1007/s00439-021-02308-w (2021).
936 46 Claussnitzer, M. *et al.* FTO Obesity Variant Circuitry and Adipocyte Browning in Humans.
937 *New England Journal of Medicine* **373**, 895-907, doi:10.1056/NEJMoa1502214 (2015).
938 47 Li, A. *et al.* mBAT-combo: A more powerful test to detect gene-trait associations from
939 GWAS data. *The American Journal of Human Genetics* **110**, 30-43,
940 doi:https://doi.org/10.1016/j.ajhg.2022.12.006 (2023).
941 48 Li, S. *et al.* The schizophrenia-associated missense variant rs13107325 regulates
942 dendritic spine density. *Transl Psychiatry* **12**, 361, doi:10.1038/s41398-022-02137-z
943 (2022).
944 49 Singh, T. *et al.* Rare coding variants in ten genes confer substantial risk for schizophrenia.
945 *Nature* **604**, 509-516, doi:10.1038/s41586-022-04556-w (2022).
946 50 Dai, Z.-M. *et al.* SECISBP2L-Mediated Selenoprotein Synthesis Is Essential for Autonomous
947 Regulation of Oligodendrocyte Differentiation. *The Journal of Neuroscience* **42**, 5860-5869,
948 doi:10.1523/jneurosci.2141-21.2022 (2022).
949 51 Sazonovs, A. *et al.* Large-scale sequencing identifies multiple genes and rare variants
950 associated with Crohn's disease susceptibility. *Nature Genetics* **54**, 1275-1283,
951 doi:10.1038/s41588-022-01156-2 (2022).
952 52 Hutchinson, A., Watson, H. & Wallace, C. Improving the coverage of credible sets in
953 Bayesian genetic fine-mapping. *PLOS Computational Biology* **16**, e1007829,
954 doi:10.1371/journal.pcbi.1007829 (2020).
955 53 Samaddar, A., Maiti, T. & de los Campos, G. Bayesian Hierarchical Hypothesis Testing in
956 Large-Scale Genome-Wide Association Analysis. *bioRxiv*, 2024.2002.2026.582204,
957 doi:10.1101/2024.02.26.582204 (2024).
958 54 O'Connor, L. J. *et al.* Extreme Polygenicity of Complex Traits Is Explained by Negative
959 Selection. *The American Journal of Human Genetics* **105**, 456-476,
960 doi:10.1016/j.ajhg.2019.07.003 (2019).
961 55 Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex
962 trait gene targets. *Nature Genetics* **48**, 481-487, doi:10.1038/ng.3538 (2016).

963
964

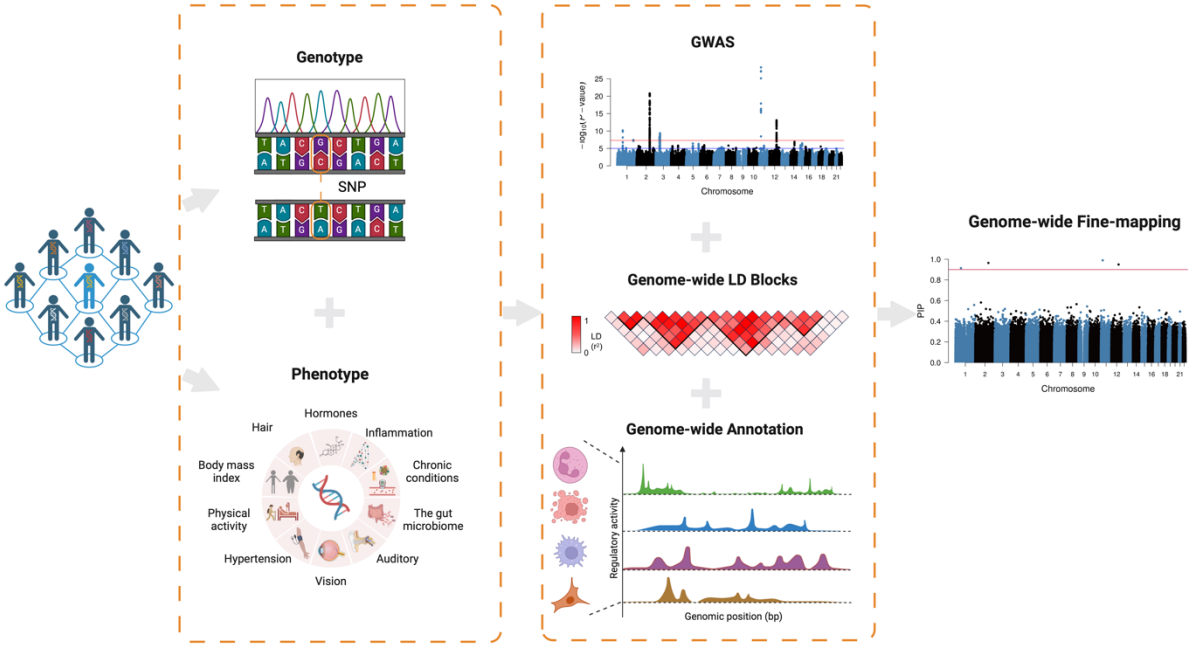

Figure 1 Schematic overview of genome-wide fine-mapping analysis using GBMM. GBMM requires the GWAS summary statistics and genome-wide LD reference to fine-map the likely causal variants for complex traits, and can incorporate functional annotations. Compared to regional-based fine-mapping approaches, GBMM estimates priors with genome-wide SNPs and MCMC sampling algorithm, and is more flexible on the assumption of the underlying distribution of causal effects (**Table S1**). The illustration was created with BioRender.com.

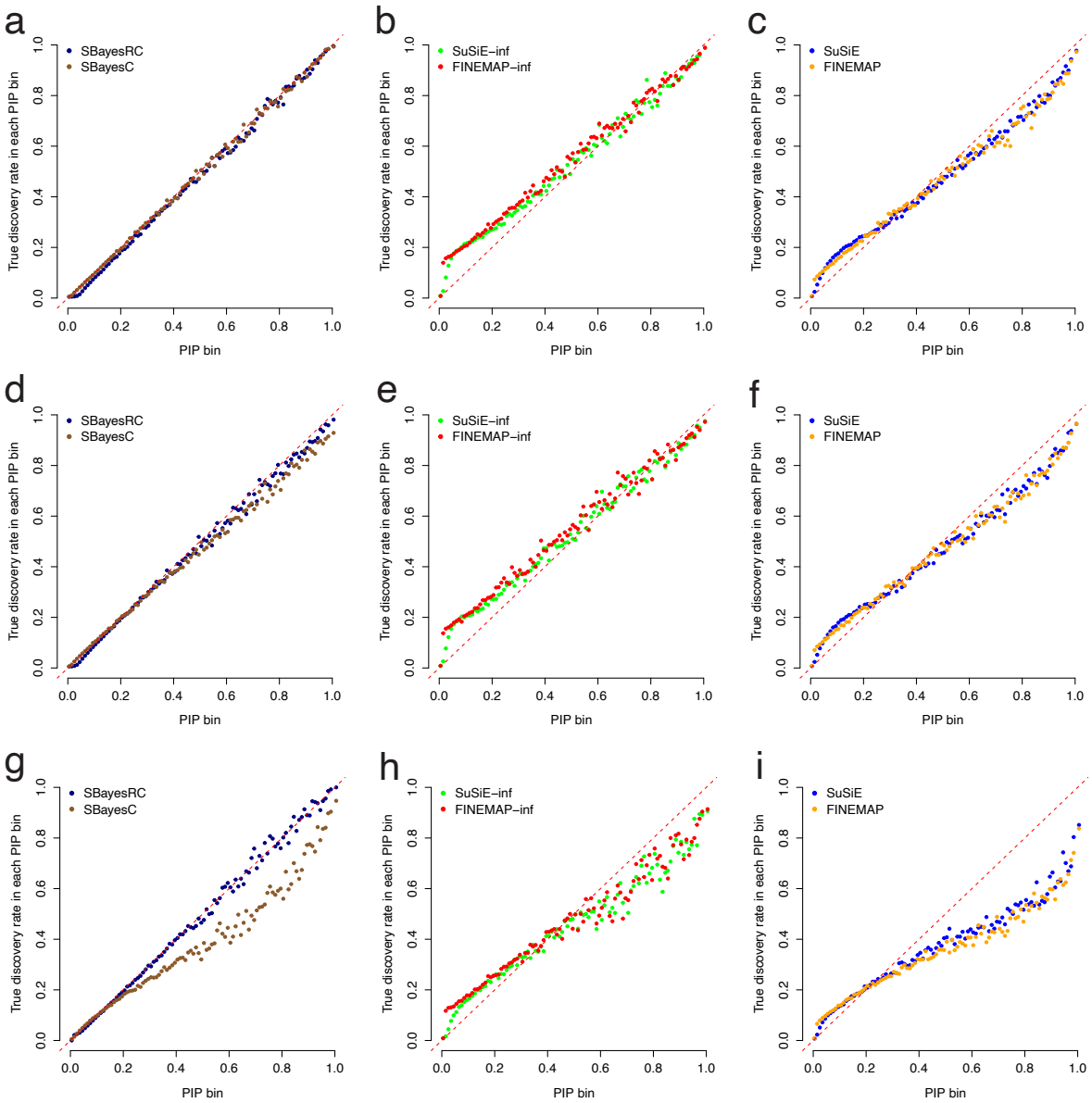

Figure 2 Comparison in the calibration of PIP between GBMM and existing fine-mapping methods under simulations with various genetic architectures. Shown are relationship between PIP and the true discovery rate across 100 PIP bins. Results showed in each column correspond to the results from GBMM (SBayesC and SBayesRC), SuSiE-inf and FINEMAP-inf and SuSiE and FINEMAMP respectively. Results shown in each row correspond to the sparse genetic architecture, major gene genetic architecture and LDMS architecture respectively.

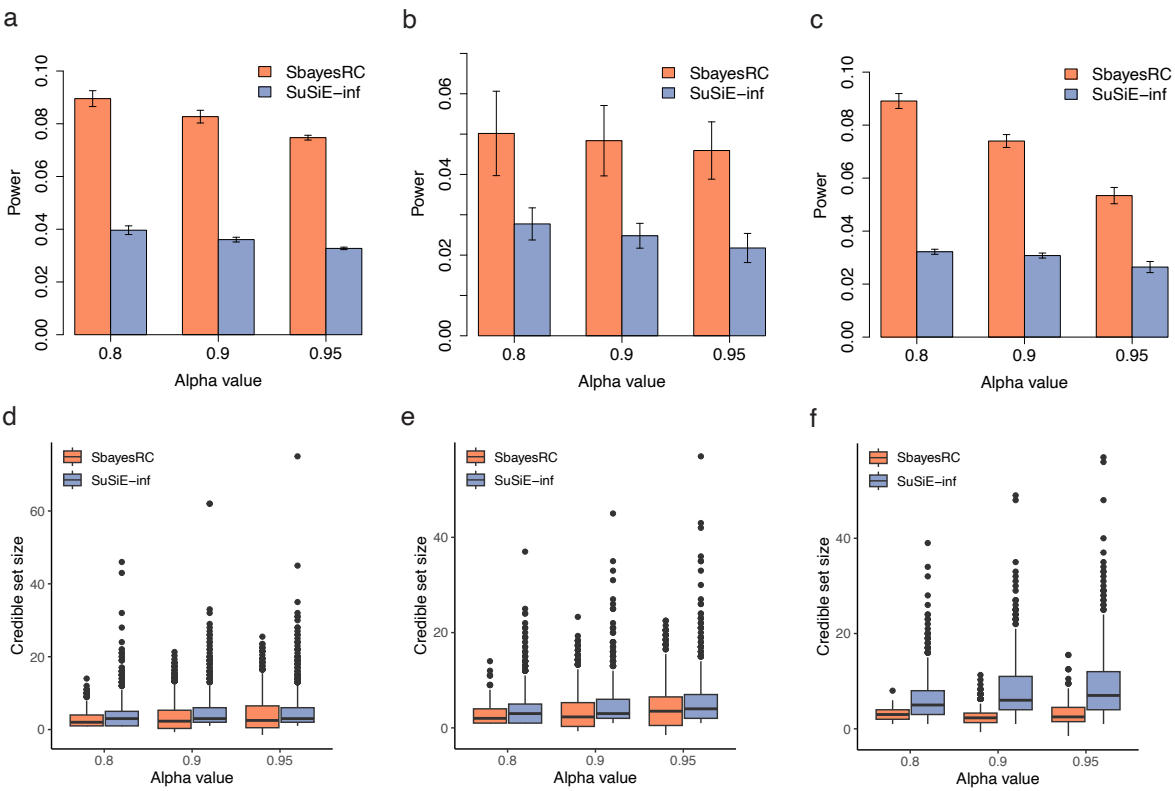

Figure 3 Comparison in local credible set (LCS) between SBayesRC and SuSiE-inf. Shown in panels (a-c) are power comparison between SBayesRC and SuSiE-inf at the same alpha cutoff. Shown in panels (d-f) are credible size comparison between SBayesRC and SuSiE-inf at the same alpha cutoff. Results showed in each column correspond to the simulation under sparse model (a, and d), major gene model (b and e) and LDMS model (c and f).

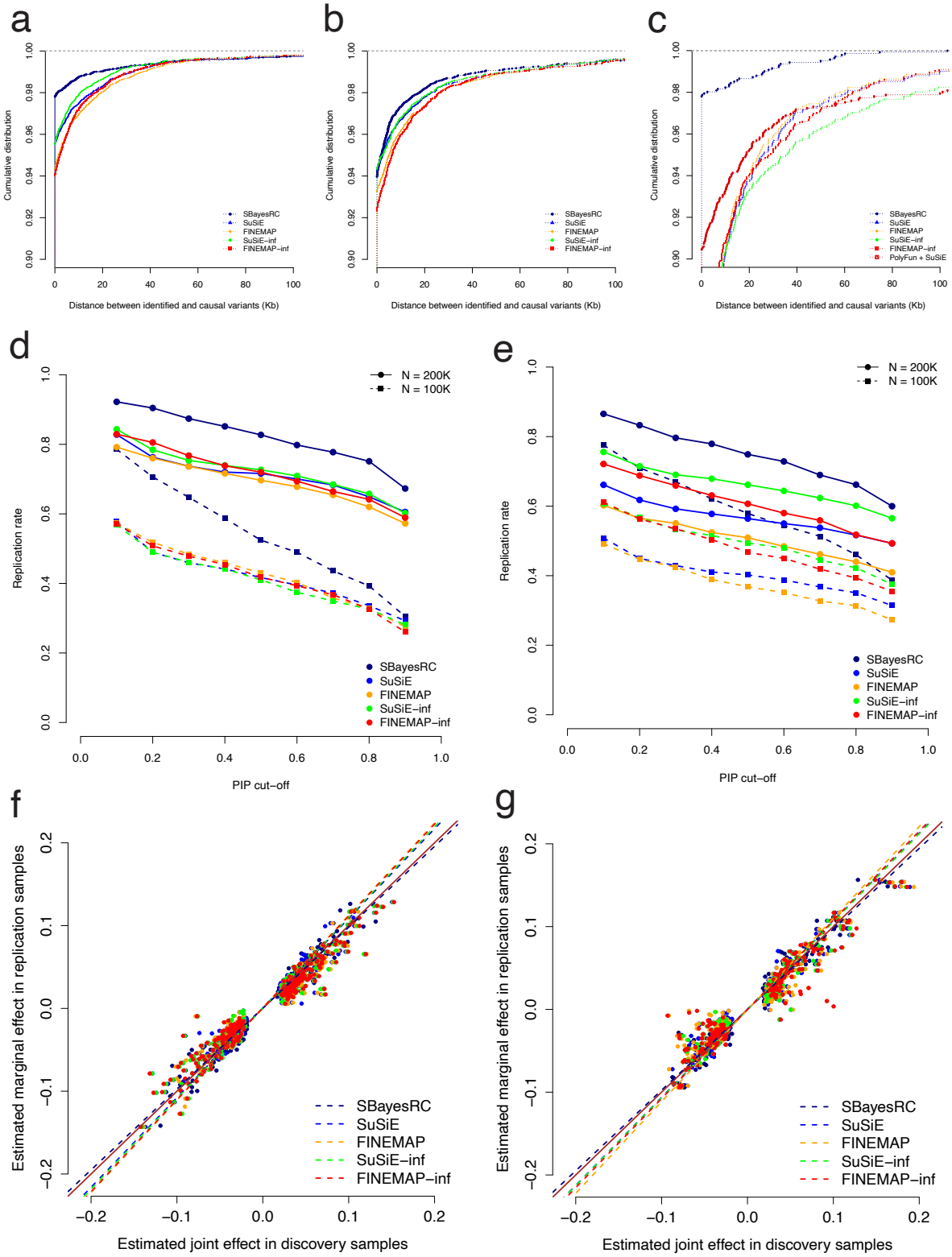

Figure 4 Comparisons of mapping precision, independent sample replication, and effect size estimation bias among fine-mapping methods. Panel (a-c) shows the distance between the causal variants and the SNPs identified by different methods at PIP of 0.9 in simulations based on sparse (a), large effects (b) and LDMS (c) genetic architectures (**Methods**). Panel (d-e) show the replication rate of discovery using different methods at a given PIP threshold in the replication

sample (x-axis) using simulations (d) and real data analysis for height in the UKB (e). Simulations are based on a sparse model and results in (d) are the mean values over ten simulation replicates. Panel (f-g) show the regression of the estimated marginal effect size in replication samples on the estimated joint effect size in discovery samples using different fine-mapping methods. Dash line shows the regression slope, which is closer to one for a less biased method. The marginal effect estimated in the independent replication samples was used as a proxy to the true value because it is an unbiased estimate. The brown solid line is the y=x line.

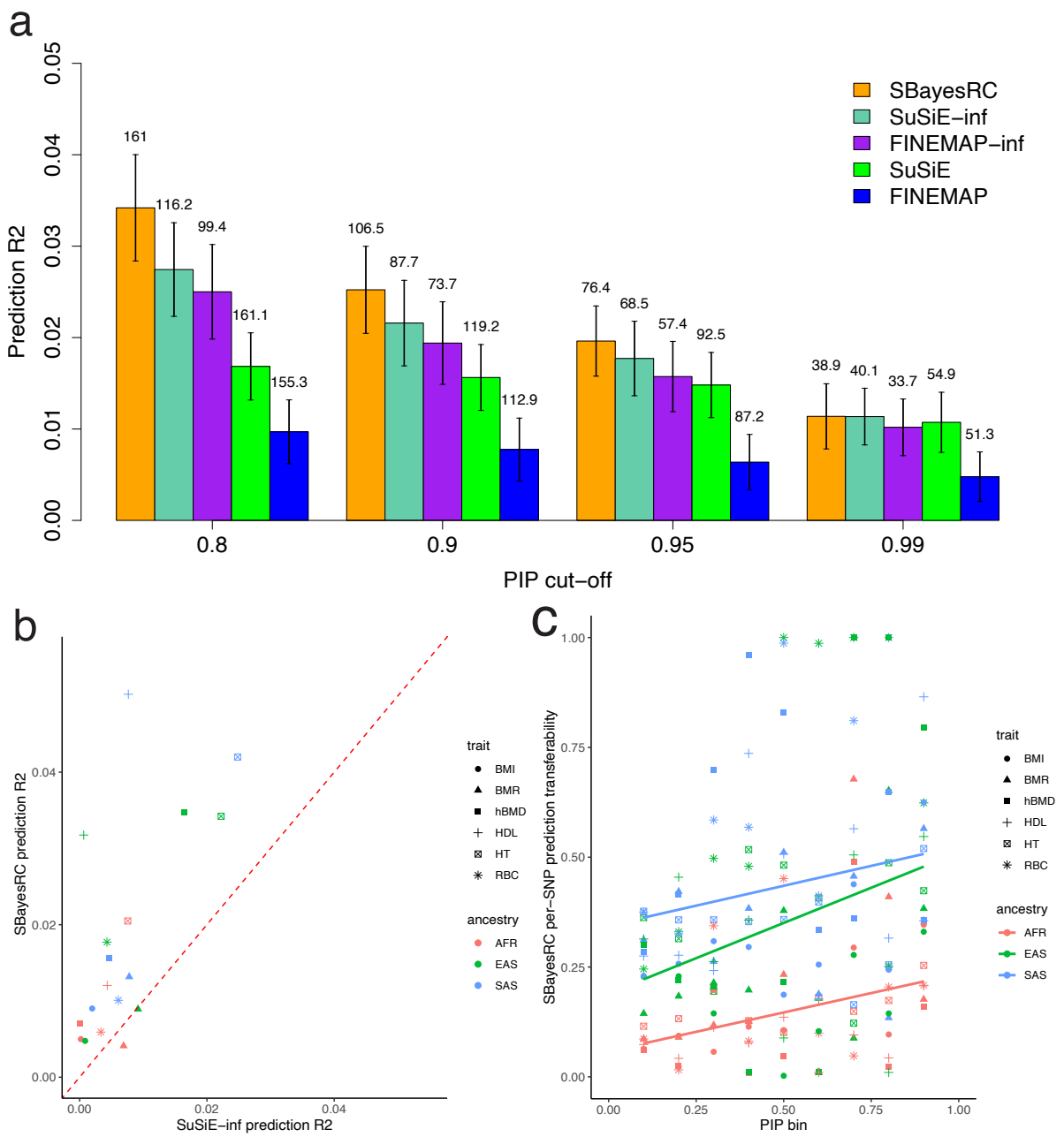

Figure 5 Out-of-sample prediction accuracy using identified variants from different fine-mapping methods. (a) Comparison of the prediction R^2 using the fine-mapped SNPs from different methods in the simulation based on the sparse architecture (**Methods**). The number above each bar is the number of fine-mapped SNPs from each method at different PIP cut-offs. (b) Comparison of trans-ancestry prediction accuracy using fine-mapped variants from SBayesRC and SuSiE-inf from the analysis of samples of European ancestry for 6 complex traits in the UK Biobank, with variants with PIP > 0.9. (c) The relationship between trans-ancestry prediction transferability and PIP in European ancestry. The transferability was computed as non-EUR-R^2/ EUR-R^2. The solid lines are regression lines across traits in each ancestry. Results are the mean values over 100 simulation replicates.

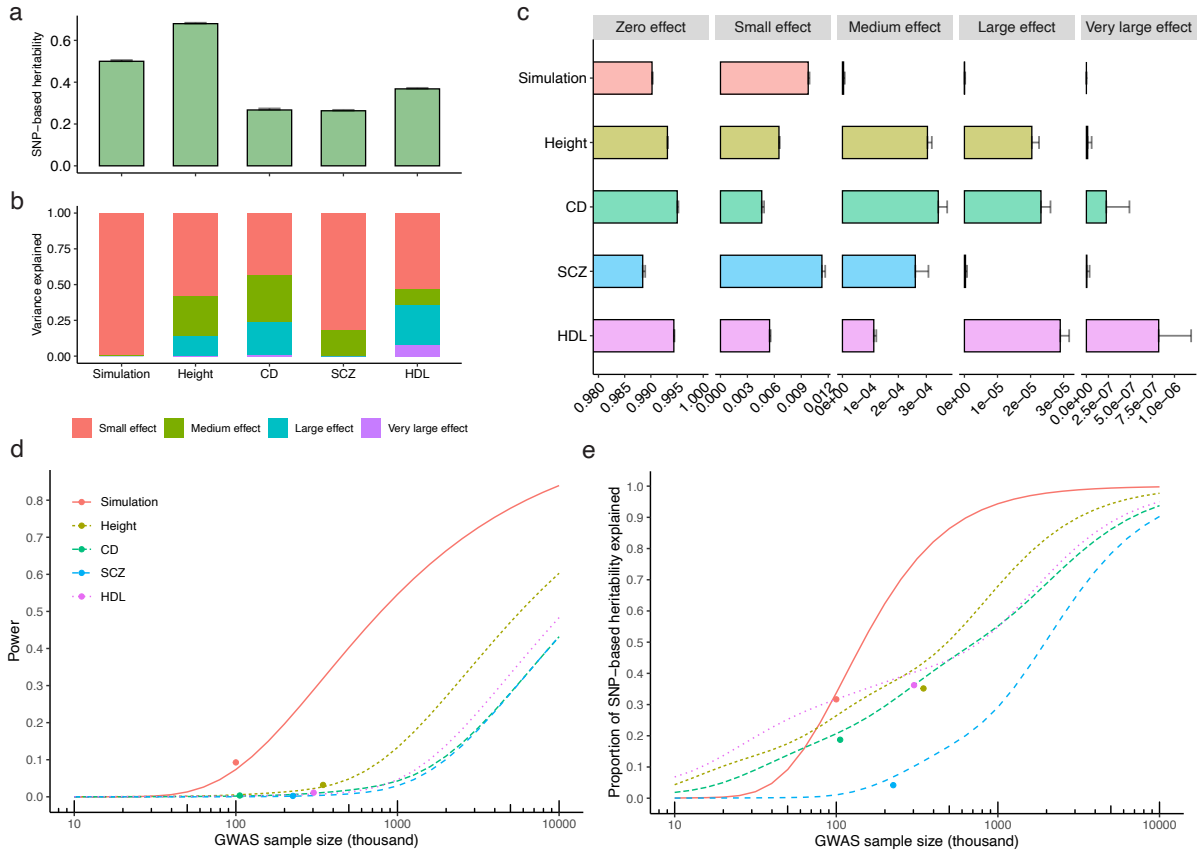

Figure 6 Projection of genome-wide fine-mapping outcomes to the theoretical power prediction in complex traits with diverse genetic architectures. (a-c) show the results of SBayesRC estimation for the SNP-based heritability (h^2_{SNP}) (a), the proportions of h^2_{SNP} explained by different mixture components (b), and the proportions of SNPs with effects from different mixture components ($\boldsymbol{\pi}$) (c), for the simulated trait, height, Crohn's disease (CD), schizophrenia (SCZ), and high density lipoprotein (HDL). (d-e) shows the theoretical prediction of the power of identifying causal variants (d) and the proportion of h^2_{SNP} explained by the identified causal variants at different GWAS sample sizes for these traits. Dot shows the observed trait outcome based on local credible sets (including singleton LCSs) identified from SBayesRC. Note that the results shown in (a, c) were used as input data for our method that predicts fine-mapping power given sample sizes (d-e).

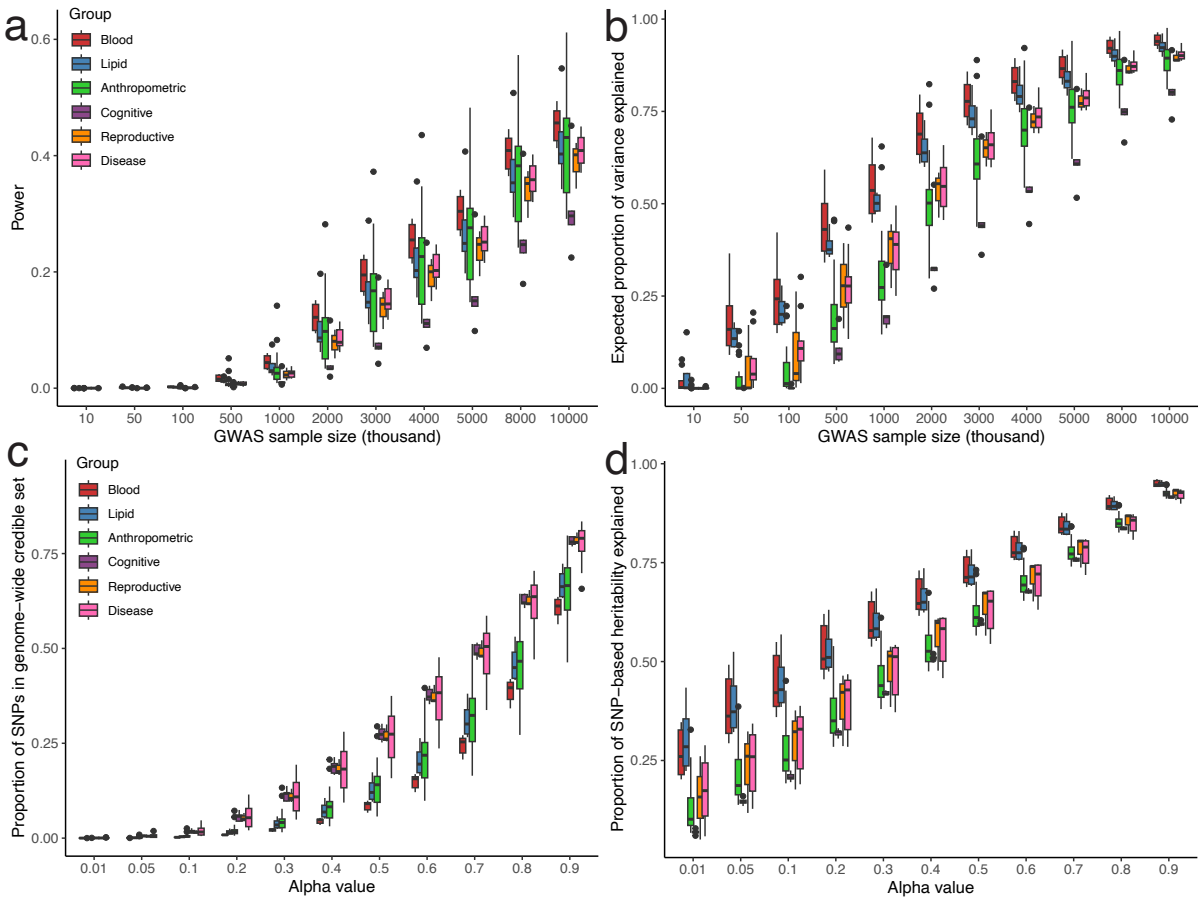

Figure 7 Theoretical identification and prediction of genome-wide credible SNPs across 48 independent complex traits. Panel (a-b) shows the theoretical prediction of power and proportion of SNP-based heritability explained by GCS SNPs at different GWAS sample sizes for the 48 complex traits, respectively. Panel (c) shows the proportion of identified GCS SNPs at different alpha threshold (proportion of causal variants captured) for the 48 complex traits (average sample size = 291K). Panel (d) shows the proportion of h^2_{SNP} explained by the GCS SNPs at different alpha threshold. Colours indicate different trait categories.

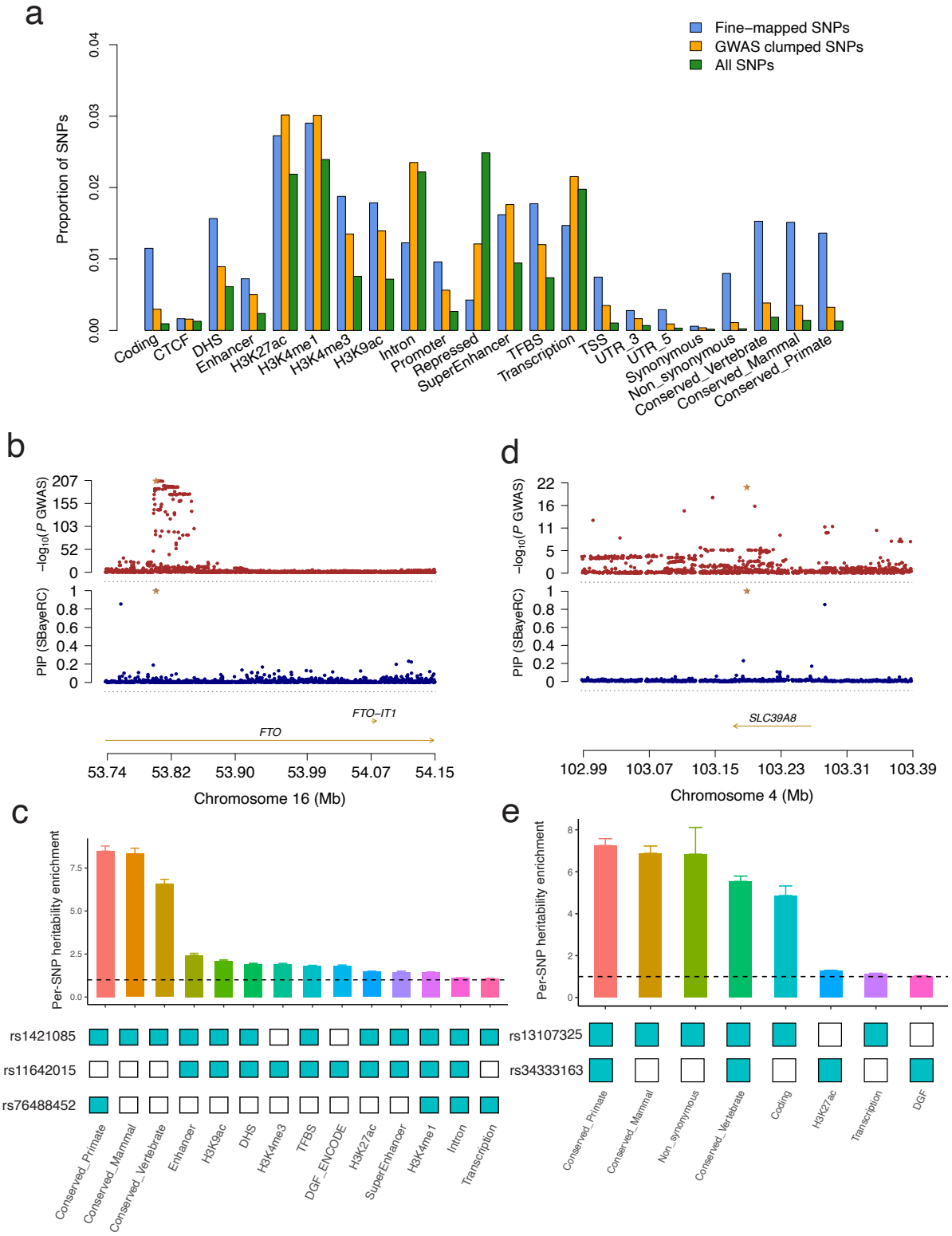

Figure 8 Genome-wide fine-mapping with functional annotations helped pinpoint the putative causal variants. Panel (a) shows enrichment of the genome-wide fine-mapped SNPs from SBayesRC and GWAS clumped SNPs in the 22 main functional categories defined in the LDSC baseline model. Panel (b) shows the prioritized causal variant at the *FTO* locus for BMI. The top track shows the *FTO* locus plot of the standard GWAS for BMI, and the second track shows the

similar plot but with the PIP from SBayesRC for BMI. The starred SNP is the known causal variant (rs1421085) for obesity at the *FTO* locus supported by previous functional studies. Panel (c) shows the per-SNP heritability enrichment for the causal variant (rs1421085), the GWAS lead variant (rs11642015) and the secondary signal (rs76488452) for BMI at the *FTO* locus. The annotations on the x-axis were those present at least once in these three variants, excluding quantitative annotations and duplicated annotations with flanking windows. Panel (d) shows the prioritized causal variant at the *SLC39A8* locus for SCZ. The top track shows the *SLC39A8* locus plot of the standard GWAS for SCZ, and the second track shows the similar plot but with the PIP from SBayesRC for SCZ. The starred SNP is the missense variant (rs13107325) fine-mapped for SCZ at the *SLC39A8* locus. Panel (e) shows the per-SNP heritability enrichment for the causal variant (rs13107325) and the secondary signal (rs34333163) for SCZ at the *SLC39A8* locus.