CARMA: Novel Bayesian model for fine-mapping in meta-analysis studies

Zikun Yang[1], Chen Wang[1,2], Linxi Liu[3], Atlas Khan[2], Annie Lee[4], Badri Vardarajan[4], Richard Mayeux[4], Krzysztof Kiryluk[2], Iuliana Ionita-Laza[1,#]

[1] Department of Biostatistics, Columbia University, New York

[2] Division of Nephrology, Department of Medicine, College of Physicians and Surgeons, Columbia University, New York

[3] Department of Statistics, University of Pittsburgh, Pittsburgh, PA

[4] Department of Neurology, College of Physicians and Surgeons, Columbia University, New York

[#]Correspondance: ii2135@columbia.edu

Abstract

We propose a novel Bayesian model for fine-mapping in order to identify putative causal variants at GWAS loci. Relative to existing fine-mapping methods, the proposed model has several appealing features, such as flexible specification of the prior distribution of effect sizes, joint modeling of summary statistics and large number of functional annotations, and accounting for discrepancies between summary statistics and external linkage disequilibrium values in meta-analysis settings. Using simulations, we compare performance with commonly used fine-mapping methods, including SuSiE and fastPAINTOR, and show that the proposed model has higher power and lower FDR when including functional annotations, and higher power, lower FDR and higher coverage for credible sets in meta-analysis settings. We further illustrate our approach by applying it to a meta-analysis of Alzheimer's Disease GWAS data where we prioritize putatively causal variants and genes, including *TREM2*, *ABI3*, *CASS4*, *SORL1*, *APH1B*, *BIN1*, *KAT8*, *ABCA7*, *INPP5D*, *SLC24A4*, *NDUFS2*, *MPO*, *PTK2B*, *EPHA1*, *CD33*, *RABEP1*, *ADAM10*, *CD2AP*, *PICALM*, *PILRB*, and *CLNK*. Finally, our model can be used more generally in conjunction with other fine-mapping methods such as SuSiE to minimize false-positive findings in meta-analyses.

Introduction

Meta-analyses of GWAS studies have identified a large number of significant loci. Fine-mapping is the natural next step in order to identify putative causal genetic variants at these loci. Meta-analyses however pose several challenges that can invalidate the results from existing fine-mapping methods as also pointed out in.[1] For example, using linkage disequilibrium (LD) from external panels can create inconsistencies with GWAS summary statistics which can lead fine-mapping methods to prioritize non-causal variants.[2] Similarly, uneven sample size coverage at different variants can lead to biased posterior inclusion probability (PIP) values. To illustrate this point we show the example of a GWAS locus *SCIMP* (SLP adaptor and CSK interacting membrane protein) for Alzheimer's disease (AD). We use summary statistics from a large meta-analysis GWAS of clinically diagnosed AD and AD-by-proxy with 71,880 cases and 383,378 controls of European ancestry from three consortia.[3] The LD matrix is estimated using individuals of European descent in UK Biobank (UKBB).[4] Two fine-mapping models, SuSiE[5] and fastPAINTOR,[6] prioritize variants with low Z-scores due to discrepancies between summary statistics and LD values, whereas the results of the proposed model appear as expected (Figure 1(a-c)). In this paper we introduce a new fine-mapping method that improves the power and reduces false positives in such situations. Moreover, the proposed method can be used to improve results from existing fine-mapping methods, such as SuSiE (Figure 1(d)).

There are many statistical fine-mapping methods in the literature.[5–13] Most of the existing methods can work with summary statistics and LD information from relevant reference panels, and make certain assumptions on the number of causal variants, i.e. some methods restrict the number of possible causal variants,[7,8,11,13] while others relax this assumption by introducing a prior distribution on model space, and implementing stochastic

algorithms such as Markov Chain Monte Carlo (MCMC) to reduce the computational cost.[6,9,10] Multiplicity control is another important aspect of any Bayesian fine-mapping method in order to control the false discovery rate in the context of multiple testing,[14] and several existing methods address this issue formally by introducing prior probabilities on model space.[9,10]

Here we propose a new Bayesian model, CARMA (CAusal Robust mapping method in Meta-Analysis studies), that attempts to improve upon existing methods especially in complex meta-analysis settings as described above. Our proposed model has several technical innovations: (1) it allows in addition to the usual Normal-Gamma prior family for the effect size distribution the option of a heavy-tail Cauchy distribution, which leads to a prior distribution of effect sizes that is more adaptive to the possibly different signal-to-noise ratios across loci; (2) it jointly models summary statistics and high-dimensional functional annotations - this is different from other recent models such as PolyFun[12] that first estimate prior causal probabilities based on functional annotations and summary statistics, and then apply fine-mapping methods such as SuSiE with these prior probabilities; (3) it introduces a novel Bayesian hypothesis testing approach to account for discrepancies between summary statistics and LD from external reference panels in order to avoid an increase in false positives. We illustrate the proposed method using simulations and applications to an AD GWAS meta-analysis.

Results

Overview of the proposed model

For a given locus, we assume a standard linear model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon} \sim \text{MVN}(0, \sigma_y^2 I_n)$, where $\boldsymbol{y}$ is a $n \times 1$ standardized vector of quantitative phenotype values, $\boldsymbol{X}$ is a standardized $n \times p$ genotype matrix, $\boldsymbol{\beta}$ is a p-dimensional vector of effect sizes of variants, and $\boldsymbol{\epsilon}$ is a Gaussian noise vector. Let $\boldsymbol{Z} = \left(n\sigma_y^2\right)^{-\frac{1}{2}} \boldsymbol{X}'\boldsymbol{y}$ denote the vector of single-SNP Z-scores, then $\mathbf{E}\left[\boldsymbol{Z}|\boldsymbol{\beta}, \boldsymbol{X}\right] = \left(n\sigma_y^2\right)^{-\frac{1}{2}} \boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}$ and $\text{Var}\left[\boldsymbol{Z}|\boldsymbol{\beta}, \boldsymbol{X}\right] = \boldsymbol{X}'\boldsymbol{X}/n$. We propose a method that can work with summary statistics, i.e. single-SNP Z-scores from a standard GWAS analysis, and LD from an external reference panel as previously done in GCTA,[15] fastPAINTOR,[6] and FINEMAP.[10] Let $\boldsymbol{\Sigma} = \boldsymbol{X}'\boldsymbol{X}/n$ denote the LD matrix, and $\boldsymbol{\lambda} = \sqrt{n}\boldsymbol{\beta}$. Assuming $\sigma_y^2 = 1$, then the sampling distribution of $\boldsymbol{Z}$ can be written as:

$$\boldsymbol{Z}|\boldsymbol{\lambda}, \boldsymbol{\Sigma} \quad \sim \quad \text{MVN}(\boldsymbol{\Sigma}\boldsymbol{\lambda}, \boldsymbol{\Sigma}).$$

Let $\boldsymbol{\gamma}' = \{0,1\}^p$ denote an indicator vector, such that $\gamma_i = 1$ iff $\lambda_i \neq 0$ ($\beta_i \neq 0$). Given any $\boldsymbol{\gamma}$, we assume a spike-and-slab prior:

$$\boldsymbol{\lambda}_{\boldsymbol{\gamma}}|\tau, \boldsymbol{\gamma} \quad \sim \quad \text{MVN}(0, \frac{1}{\tau} I_{\boldsymbol{\gamma}}),$$

where $\boldsymbol{\lambda}_{\boldsymbol{\gamma}}$ represents the vector of non-zero entries of $\boldsymbol{\lambda}$, i.e. $\gamma_i = 1$. The value of the prior precision parameter τ may be tuned using empirical evidence (i.e. with 95% probability, the causal variants at a locus explain less than 1% of the trait variation). We want to identify the true model that generated the summary statistics through posterior inference within a Bayesian paradigm.

We note that the precision parameter τ can be assigned a prior distribution to yield a mixture of normal distributions on $\boldsymbol{\lambda}$. This option provides more flexibility and robustness in situations when tuning τ to be consistent with the underlying model at a given locus may be difficult (e.g. for loci very different effects). Specifically, integrating out τ against a prior distribution $\text{Gamma}(0.5, 0.5)$ yields a heavy-tailed Cauchy prior distribution on $\boldsymbol{\lambda}$, which provides an adaptive Bayesian inference under different signal-to-noise ratios. Although we focus our description in the main text on the slab-and-spike prior, we derive the corresponding results for the Cauchy prior distribution in the Supplemental Material, and show simulation results (both the spike-and-slab and Cauchy priors are provided as options in the software package implementing CARMA).

We additionally assume a truncated Poisson prior on the size of the model, i.e. $\sum_{i=1}^{p} \gamma_i \sim \text{Truncated Poisson}\,(\eta)$ and $\sum_{i=1}^{p} \gamma_i \in \{0, 1, \ldots, p\}$, to control the total number of causal variants assumed by a given model, and hence provide multiplicity control (or control of the false discovery rate). We implement a Shotgun stochastic search

algorithm[16] for sampling from the posterior distribution over the model space, which has advantages over the more commonly used MCMC-based methods in that it leads to a more complete exploration of the areas with high posterior probability in the model space. This semi-exhaustive search feature alleviates the problem of unequal PIPs for perfectly correlated variants that can happen with other MCMC algorithms.

High-dimensional functional annotations. CARMA seeks to distinguish among highly correlated variants by leveraging functional annotations. Let $\boldsymbol{W}$ be a $p \times (q+1)$ matrix for p SNPs and q functional annotations and a p-dimensional $\mathbf{1}$ vector. We maximize the joint likelihood of the summary statistics and functional annotations

$$L(\boldsymbol{\theta}; \boldsymbol{Z}, \boldsymbol{\gamma}, \boldsymbol{W}) = m(\boldsymbol{Z}|\boldsymbol{\gamma}) \Pr(\boldsymbol{\gamma}|\boldsymbol{W}, \boldsymbol{\theta}),$$

where $\boldsymbol{\theta}$ is a $(q+1)$-dimensional vector of the corresponding coefficients, through an expectation–maximization (EM) algorithm with the prior probability $\Pr(\boldsymbol{\gamma}|\boldsymbol{W}, \boldsymbol{\theta})$ being linked to functional annotations via a penalized logistic model.[17]

Bayesian hypothesis testing for meta-analysis settings. A principal advantage of the proposed model over existing fine-mapping models is in the context of meta-analyses when discrepancies between GWAS summary statistics and LD from external reference panels can lead to false positive results when using existing methods as also noted in.[1] A first type of discrepancy is due to the external LD panel itself; a second type is due to sample size heterogeneity across SNPs in the meta-analysis. To account for these possible discrepancies, we propose a Bayesian hypothesis testing procedure for detecting and removing inconsistencies, leading to a reduction in false discoveries relative to existing fine-mapping models. We also show that this procedure can be combined with genotype imputation to reduce the effect of the second type of discrepancy due to sample size inconsistencies across SNPs. More details are available in the Methods section.

Credible sets and credible models. In[5] the authors define a credible set as the smallest subset of correlated variants (with correlation within the set greater than some threshold r) that has probability ρ or greater of containing at least one causal variant, e.g. the ith and jth SNPs can form a credible set if $\Pr(\gamma_i = 1|\boldsymbol{Z}) + \Pr(\gamma_j = 1|\boldsymbol{Z}) \geq \rho$ and $\text{cor}(Z_i, Z_j) \geq r$. The credible set provides a small set of SNPs for follow-up studies.

However, as also mentioned in,[5] a SNP should not be considered as non-causal if it is not included in any credible set. For example, there are circumstances where a group of highly correlated SNPs containing one causal SNP might not result in a sufficiently large PIP to form a credible set. Hence, we introduce the concept of credible model at a locus, as a complement to credible sets, based on the top candidate models in terms of posterior probability. This is similar to FINEMAP[10] which provides configuration-specific posterior probabilities. Let $\boldsymbol{\gamma}_{(1)}$ denote the leading model that receives the largest posterior probability among the visited candidate models; then we define as credible model at a locus the set of candidate models such that the posterior odds between the leading model and the models in the set is smaller than a pre-determined threshold, such as 3.2 or 10 as recommended in.[18] We view the credible model and the variants identified by a credible model as complementary to credible sets. Compared to credible sets, credible models typically involve less variants while identifying a higher proportion of causal variants as shown in simulations below.

Simulations

We perform simulations to investigate the performance of CARMA and competitor methods, fastPAINTOR and SuSiE, first in a homogeneous setting with in-sample LD and then in more realistic heterogeneous settings for meta-analyses with LD estimated from external reference panels.

Generating simulated datasets (in-sample LD)

Genotype simulation. We use the R package 'sim1000G'[19] to simulate genotypes based on the 1000 Genomes Project data (phase 3, European population). To select regions representative of GWAS loci in

terms of size and LD structure, we focus on 94 loci identified as risk regions in a recent GWAS on breast cancer.[20] Within each region, we filter out rare variants (MAF<0.01); the number of variants in each region ranges between $\sim 1,500-4,000$. We simulate genotype data for $n = 10,000$ individuals.

Causal SNP selection. To generate prior causal probabilities we use data on 200 DeepSEA chromatin features.[21] Let $\boldsymbol{\theta}$ denote a 201-dimensional coefficient vector including the intercept term; we randomly select 100 chromatin features as being related to the causal status of variants, with the coefficients being sampled from $N(0, 0.1^2)$. Hence the impact of each individual annotation is weak. We used a simulation setting similar to that in PolyFun, i.e. the prior causal probability for each SNP is proportional to a linear combination of annotations, such as $\Pr(\gamma_i = 1) \propto \boldsymbol{w}_i\boldsymbol{\theta}$, where $\boldsymbol{w}_i$ is the vector of annotations. The causal SNPs are selected among those with highest prior probabilities in such a way that the absolute value for the pairwise Pearson correlation between any two causal variants is at most 0.1. For a given locus, let T denote the index set of the true causal SNPs, i.e., if $i \in T$, then $\gamma_i = 1$. In simulations we assume $|T| \in \{1, 2, 3\}$ at each locus (results for $|T| = 2$ are shown in the main manuscript; results for $|T| \in \{1, 3\}$ are shown in Supplemental Figures).

Phenotype generation. For each $i \in T$, the effect sizes of the causal SNPs are drawn independently from $\beta_i \sim N(0, 0.5^2)$, and for all $i \notin T$ we set $\beta_i = 0$. The phenotypic variance σ_y^2 is computed such that $\phi = 0.0075$ at each locus, where $\phi = \frac{\mathrm{Var}(\boldsymbol{X}\boldsymbol{\beta})}{\sigma_y^2+\mathrm{Var}(\boldsymbol{X}\boldsymbol{\beta})}$. Then we sample $\boldsymbol{y}$ such that $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$; $\boldsymbol{\epsilon} \sim \mathrm{N}(0, \sigma_y^2 I_{n\times n})$.

Remarks on some implementation details:

1. We assume two scenarios (1) no functional annotation and (2) with functional annotations (200 DeepSEA chromatin features[21]). SuSiE can only include one annotation in the form of prior probabilities. We estimated PolyFun prior probabilities using the PolyFun package in R, then we provided the estimated prior probabilities to SuSiE. fastPAINTOR cannot handle high-dimensional annotations therefore we select a subset of the ten most informative annotations as follows: we first compute correlations between the true linear predictor $\boldsymbol{w}_i\boldsymbol{\theta}$ and each functional annotation, then we iteratively select top correlated annotations such that no annotation has absolute correlation > 0.3 with previously selected annotations. Then, fastPAINTOR is run based on the selected annotations.

2. For CARMA, the value of the hyper-parameter η is set at the default setting 1 throughout the simulations. The prior precision parameter τ (in the case of the spike-and-slab prior) is set at 0.0512 corresponding to an assumed explained variance at the locus $\phi = 0.0075$ and sample size $n = 10,000$. More simulation studies for different values of η and τ can be found in the Supplemental Material (Figures SI2, SI3, and SI5).

3. We run fastPAINTOR and CARMA models at a chromosome level, while SuSiE is run one locus at a time (since SuSiE does not have the option to aggregate multiple loci).

4. We run each model with the default number of maximum causal variants per locus assumed by each model, i.e. CARMA and SuSiE assume 10 causal variants and fastPAINTOR assumes 2 causal variants per locus.

Credible sets. For SuSiE, we use the credible sets reported by SuSiE. For CARMA and fastPAINTOR we compute credible sets as in[5] (Methods). We assess different measures of performance with respect to credible sets from each model, including:

Power: The overall proportion of simulated causal variants included in any credible set.

Coverage: The proportion of credible sets that contain a causal variant.

Size: The number of variants included in a credible set.

Purity: The mean squared correlation of variants in a credible set.

Results for $\rho = 0.99$ are shown in Figure 2(a) (see Figure S1 for $\rho = 0.95$). Without annotations, CARMA and SuSiE have similar performance. In contrast, fastPAINTOR has lower power, coverage and purity. The substantially lower coverage of the credible sets from fastPAINTOR suggest the need of including penalization of the size of the selected model to control the false positives. When functional annotations are included, the power of CARMA increases relative to the scenario with no functional annotation, and relative to the competitor models, SuSiE and fastPAINTOR. These results suggest that the inclusion of functional annotations has a greater impact for the models that jointly integrate summary statistics and functional annotations (such as CARMA and fastPAINTOR) compared to SuSiE+PolyFun which is based on two separate modeling stages. Therefore algorithms that maximize the marginal likelihood conditional on both summary statistics and annotations as proposed here may be preferable.

Credible models. Credible models provide complementary information to credible sets. Therefore, we compare results based on the variants included in the credible model in CARMA and the variants included in all the credible sets for CARMA, SuSiE, and fastPAINTOR at a locus. Compared to credible sets, credible models generally achieve higher power with smaller sets of selected variants, which may facilitate follow-up investigations (Figure 2(b)).

Power and FDR based on positive predictions. We also compare methods in terms of power and false discovery rate (FDR) using positive predictions, i.e. PIP is greater than a given threshold. With no functional annotation, the performance of CARMA and SuSiE are comparable whereas fastPAINTOR has higher FDR and lower power (Figure 2(c)). With functional annotations, CARMA performs best with higher power and lower FDR. As also noted above, there is greater improvement in performance (higher power and lower FDR) for CARMA and fastPAINTOR when introducing functional annotations relative to SuSiE.

PIPs for perfectly correlated SNPs. Methods based on stochastic algorithms such as MCMC often fail to evenly explore the area of the posterior model space which creates problems when there are perfectly correlated SNPs. Due to the semi-exhaustive search feature of the Shotgun algorithm, CARMA examines the neighborhood of the currently selected model, including candidate models that exchange perfectly correlated SNPs. This way, the resulting PIPs tend to be more consistent between highly correlated SNPs. To illustrate this point, we examine the standard deviation of the PIPs within groups of perfectly correlated SNPs based on the results of the three models when no functional annotation is included. Across the 94 loci, there are 30,131 groups of perfectly correlated SNPs with an average of 317 groups at each locus, and 4.23 SNPs per group. Since the values of the PIPs returned by each model can vary across different models, we standardized the PIPs within each group by dividing the PIPs by the maximum PIP in each group. CARMA has a standard deviation of 0.06 for PIPs of perfectly correlated SNPs. This is in contrast to fastPAINTOR which has a standard deviation of 0.493. Note that SuSiE has identical PIPs for perfectly correlated SNPs by definition due to its algorithm that runs univariate regression individually.

Effect of LD-summary statistics inconsistencies in meta-analyses

In complex meta-analyses studies, with LD estimated from external reference panels, it is not uncommon to have discrepancies between Z-scores and LD values, which leads to biased PIPs for existing fine-mapping models, and false prioritization of non-causal variants as also illustrated in Figure 1. Furthermore, more complications can arise due to sample size heterogeneity across different SNPs. Allele flipping (i.e. alleles are encoded differently in the study and reference panel) can also create problems for fine-mapping.[22] Here we show the robustness of CARMA in such scenarios.

Specifically, we start with the same simulated datasets as above, i.e. 94 GWAS loci with 10,000 simulated individuals, but switch from in-sample LD to UKBB LD. We randomly separate the individuals into three non-overlapping groups with 1,000, 1,500, and 7,500 individuals. We compute summary statistics for each group, i.e. $\boldsymbol{Z}_1$, $\boldsymbol{Z}_2$, and $\boldsymbol{Z}_3$. We meta-analyze results across groups using the inverse-variance based method in METAL.[23] We first compute Z-scores using a consistent meta-analysis by meta-analyzing all three groups for

each SNP. We then consider an inconsistent meta-analysis, where the final dataset is constructed in such a way that the Z-scores of 75% of the SNPs are generated by meta-analyzing all three groups, the Z-scores for 15% of the SNPs are based on the two groups with 1,000 and 1,500 individuals, and the remaining 10% are only based on the group with the 1,000 individuals. Note that in the consistent meta-analysis setting, all Z-scores are computed based on the complete set of 10,000 individuals, and therefore the use of external UKBB LD in the only source of discrepancies. In the inconsistent meta-analysis setting, both the use of external LD and heterogeneity in sample size across SNPs leads to discrepancies between Z-scores and LD values.

Power, Coverage and FDR. CARMA's performance is significantly less affected by discrepancies between Z-scores and LD values compared with SuSiE and fastPAINTOR in terms of both power and coverage (Figure 3(a)). In the inconsistent setting, the coverage for CARMA is also reduced but still much higher than for SuSiE and fastPAINTOR (61% vs. 10% and 22%) due to sample size heterogeneity across SNPs. Using imputation before fine-mapping largely alleviates this issue, with the performance of CARMA being similar to that in the consistent setting. Without imputation, the outlier detection method will drop a significant proportion of SNPs especially due to the sample size heterogeneity across SNPs, whereas imputation greatly lowers the dropping rate (Table SI1). Similar results hold when looking at credible models (Figure 3(b)). In terms of FDR, CARMA performs substantially better than SuSiE and fastPAINTOR (Figure 3(c)). In particular, for SuSiE, there are many non-causal SNPs that are mistakenly assigned PIPs close to 1 (see below).

Credible sets with only one SNP. A particular situation arises when a credible set contains only one SNP (i.e. its PIP exceeds ρ). Such scenarios require that a SNP has a large summary statistic and be relatively independent of all the other SNPs. Discordant LD/Z-score values when using external LD may create the appearance that a SNP with large summary statistic is weakly correlated to surrounding SNPs. Figure 4(a) shows the total number of credible sets with one SNP for the three models across the 94 loci, and the corresponding coverage of these credible sets. As shown, with external UKBB LD, SuSiE has increased number of credible sets with one SNP, and corresponding poor coverage. Similarly, fastPAINTOR has poor coverage in such settings. In contrast, CARMA's performance, especially with the imputation method, appears robust.

Using CARMA and SLALOM to identify outliers. We note that recently another method, SLALOM,[1] has been described for identifying discrepancies between Z-scores and LD values in meta-analyses as discussed here and flag suspicious loci. SLALOM assumes a single causal variant per locus. For a given region with p SNPs, SLALOM utilizes the leading SNP (the highest PIP based on approximate Bayes factor (ABF) fine-mapping) as the presumed causal variant and the corresponding LD correlations to all other SNPs in the region, and conducts $p-1$ χ^2 tests for detecting suspicious loci and outlier variants at such loci. However, the assumption of only one causal variant residing in the testing region may be too strong, and may lead to low recall if there are multiple independent signals in the region.

We ran SLALOM and CARMA on the simulated datasets for the inconsistent meta-analysis scenario, and recorded the outliers identified by each method. Then, we ran SuSiE on the datasets with outliers removed by SLALOM, and CARMA respectively. Figure 4(b) show that the outlier removal by SLALOM does not make a significant impact on the performance of SuSiE. On the other hand, the performance of SuSiE has been greatly improved by removing outliers identified by CARMA. These results suggest that CARMA can successfully identify the problematic outliers and other fine-mapping methods can benefit as well.

In our analyses SLALOM only identified a small proportion of outliers, and those outliers are highly correlated to the SNP with the largest absolute Z-score, as the minimum of the absolute value of pair-wise correlations between the SNP with the largest Z-score and the outliers identified by SLALOM is ~ 0.76. Therefore, given the fact that the pairwise LD correlation between causal variants in the simulation is at most 0.1, SLALOM has missed those outliers/inconsistencies associated with secondary or other independent signals in the region besides the signal containing the leading SNP.

Fine-mapping Alzheimer's disease GWAS loci

We present fine-mapping results (primarily from CARMA) at 28 GWAS loci identified in a large meta-analysis of clinically diagnosed AD and AD-by-proxy with 71,880 cases and 383,378 controls of European ancestry.[3] The clinically diagnosed AD case-control data are from three consortia: Alzheimer's disease working group of the Psychiatric Genomics Consortium (PGC-ALZ), the International Genomics of Alzheimer's Project (IGAP), and the Alzheimer's Disease Sequencing Project (ADSP). The AD-by-proxy data are based on 376,113 individuals of European ancestry from UKBB. Note that this is a complex meta-analysis dataset, and the sample sizes available at different genetic variants can vary from 9,703 to 444,006 depending on which data are available for each SNP. We use the leading SNP at each locus from the meta-analysis (phase 3) and for the purposes of fine-mapping define the locus as ± 500kb centered around the leading SNP (Figures S7-S34). We do not include the *HLA*, *APOE* and *AC074212.3* loci due to long-range LD in these regions and extreme values of the summary statistics, which cause numerous false findings when fine-mapping these regions. We also exclude the *SUZ12P1* locus for the lack of globally significant variant ($-\log$ (p-value) ≥ 7.3) in the phase 3 data. In total we perform fine-mapping at 28 loci. We use the LD matrix from the UKBB provided by PolyFun. We also perform fine-mapping by including functional annotations provided by PolyFun (including 187 individual annotations and the PolyFun prior causal probability made available by PolyFun based on a meta-analysis of several UKBB traits[12]). Therefore, in total for CARMA we consider three scenarios: (1) no functional annotations, (2) PolyFun prior causal probability, and (3) 187 annotations provided by PolyFun plus PolyFun prior causal probability.[12]

SuSiE fine-mapping with outliers removed by CARMA. For 12 loci (*CR1*, *BIN1*, *CLNK*, *HS3ST1*, *CD2AP*, *ZCWPW1*, *MS4A6A*, *PICALM*, *ADAM10*, *SCIMP*, *ABI3*, and *ABCA7*), SuSiE reports multiple SNPs with PIPs equal or close to 1 due to discrepancies between Z-scores and LD values (Figures S35-S46), e.g. multiple credible sets with single SNPs with weak Z-scores. We have removed the outliers identified by CARMA and provided the resulting filtered data to SuSiE. By filtering out the outliers identified by CARMA, SuSiE can be properly executed (Table 1 and Figures S35-S46).

Next we focus on CARMA fine-mapping results.

Enrichment of fine-mapped SNPs in functional categories. To assess the enrichment of fine-mapped SNPs in several primary functional categories in the genome, we estimated the functional enrichment of fine-mapped SNPs in credible models for six selected binary categories (including non-synonymous, conserved, 3' UTR flanking and 5' UTR flanking, H3K4me3, repressed) where the enrichment is defined as the proportion of SNPs included in the credible models of CARMA lying in a functional category divided by the proportion of genome-wide SNPs lying in the same functional category. Even when not using annotations we see evidence for enrichment for variants in credible models for the five active functional categories (non-synonymous, conserved, 3' UTR flanking and 5' UTR flanking, H3K4me3,), and depletion in the repressed category (Figure 5). By integrating PolyFun annotations, the functional enrichments for SNPs in credible models become stronger especially for non-synonymous, conserved and 3' UTR/5' UTR flanking categories.

Fine-mapping results at 28 GWAS loci. We show fine-mapping results for individual loci using the three CARMA scenarios in Figures S47-S74. For a quick glance at the results we highlight in Table 2 those SNPs with the largest PIP in the analysis with the PolyFun annotations (even though these SNPs may not necessarily be the causal SNPs). Furthermore, we have used several functional annotation tools including CADD,[24] Eigen/EigenPC,[25] data on eQTL and sQTL from GTEx[26] and mQTL from The Religious Orders Study and Memory and Aging Project (ROSMAP)[27] to functionally annotate these SNPs. We also report for each SNP the gene with the highest V2G score in the Open Targets Genetics Portal. There is a rather high concordance between the gene with highest score in Open Targets and the gene identified by mQTLs in Brain Dorsolateral Prefrontal Cortex in ROSMAP (19/28; including *BIN1*, *SORL1*, *APH1B*, *ABCA7*, *NDUFS2*, *PTK2B*, *KAT8*, *INPP5D*, *CD2AP*, *PILRB*, *EPHA1*, *CD33*, *PICALM*, *RABEP1*, *CASS4*, *SLTM*, *CLNK*, *RABEP1* and *MPO*). Many of these top PIP SNPs are also eQTLs in brain or whole blood tissues in GTEx. Additionally, we have

identified two rare missense variants (rs616338 in *ABI3* and rs143332484 in *TREM2*) both with PIP>0.99 in the analysis including PolyFun annotations. Both of these genes are highly expressed in microglia cells and the same two variants were identified in a previous AD association study.[28] These fine-mapped SNPs and the putative effector genes could be interesting for follow-up investigations.

Discussion

We have proposed here a novel Bayesian fine-mapping method, CARMA, which is designed to prioritize potentially causal variants within GWAS risk loci by leveraging the LD structure and functional annotations available for variants at the locus under investigation. CARMA allows flexible specification of the prior distribution of effect sizes, including a spike-and-slab prior and a heavy-tail Cauchy distribution, and jointly maximizes the likelihood of summary statistics and functional annotations in a unified EM algorithm with multiplicity control. Importantly, we introduce a novel Bayesian hypothesis testing approach to account for mismatches between summary statistics and LD values from external reference panels in order to avoid prioritization of non-causal variants. Through extensive simulations, we demonstrate that CARMA has higher precision and lower FDR relative to SuSiE and fastPAINTOR, especially in complex meta-analyses settings with LD from external reference panels. Furthermore, in such settings the outlier observations identified by CARMA can be used to improve the performance of other fine-mapping methods, including SuSiE.

Our procedure for outlier detection and removal may lead to true causal variants being removed from the analysis. In particular, when summary statistics at SNPs in a region are based on different sample sizes, a possible solution would be to perform imputation before running CARMA. Although imputation is also affected by the reference LD being used, and can therefore introduce outliers on its own, we have found that it can be of great help in retaining more SNPs including potentially causal variants in the analyses.

We further illustrate the use of CARMA in a large fine-mapping analysis of 28 GWAS loci for Alzheimer's disease identified in.[3] The results show that CARMA has the ability to better handle possible discrepancies between the LD matrix from the UKBB and the summary statistics in the meta-analysis relative to competitor methods such as SuSiE. In particular, for 21/28 loci we highlight a putative effector gene that has multiple lines of evidence including from CARMA, external eQTL/mQTL databases and effector gene mapping tools such as V2G in OpenTargets.

Very recently, a new version of SuSiE has been released that is applicable to summary statistics and LD matrix extracted from reference panels.[22] In this context, the authors developed a likelihood ratio test for identifying a particular type of inconsistency, namely the "allele flip" scenario. Note that our proposed outlier detection method is more general and deals with broad types of inconsistencies, including those generated by different sample size coverage at different SNPs in the data. We also note that the diagnostic procedure in SuSiE is a separate step, i.e. not integrated into the main algorithm, and requires the inversion of the entire LD matrix for identifying one outlier, which is computationally intensive. In our applications to the motivating example in Figure 1 and other loci with biased PIPs in our analyses, this new implementation has failed to solve the problem (more details are in the Supplemental Material).

CARMA has been implemented in a computationally efficient R package.

Methods

Basic notations and assumptions

For a given locus, we assume a standard linear model:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon} \sim \text{MVN}(0, \sigma_y^2 I_n),$$

where $\boldsymbol{y}$ is a standardized $n \times 1$ vector of quantitative phenotype values, $\boldsymbol{X}$ is a standardized $n \times p$ genotype matrix, $\boldsymbol{\beta}$ is a p-dimensional vector of effect sizes of SNPs, and $\boldsymbol{\epsilon}$ is a Gaussian noise vector. GWAS are usually performed in a univariate fashion, so that for each variant i with $i = 1, \ldots, p$, we compute the marginal Z-score

$Z_i = \frac{\boldsymbol{x}_i'\boldsymbol{y}}{\sqrt{\sigma_y^2 n}}$, as in other fine-mapping methods[10].[6] Let the SNP correlation (LD) matrix be $\boldsymbol{\Sigma} = \frac{\boldsymbol{X}'\boldsymbol{X}}{n}$; note that $\boldsymbol{\Sigma}$ can be approximated using the appropriate, population-matched reference panel. Assuming $\sigma_y^2 = 1$, a common practice in fine-mapping, and $\boldsymbol{\lambda} = \sqrt{n}\boldsymbol{\beta}$ (i.e., $\lambda_i \neq 0$ if and only if $\beta_i \neq 0$), the sampling distribution of $\boldsymbol{Z}$ can be represented as

$$\boldsymbol{Z}|\boldsymbol{\lambda}, \boldsymbol{\Sigma} \sim \text{MVN}(\boldsymbol{\Sigma}\boldsymbol{\lambda}, \boldsymbol{\Sigma}).$$

Details are in the Supplemental Material.

Indicator vector $\boldsymbol{\gamma}$ and true model. Let $\boldsymbol{\gamma}' = \{0,1\}^p$ denote an indicator vector, such that $\lambda_i \neq 0$ iff $\gamma_i = 1$. Also, let $S = \{i; \lambda_i \neq 0\}$ denote the index set of causal SNPs assumed by a particular model, such that if $i \in S$, then $\gamma_i = 1$ and $\lambda_i \neq 0$. Then, each indicator vector $\boldsymbol{\gamma}_S$ uniquely defines a model with dimension equal to $|S| = \sum_{i=1}^p \gamma_i$. For a particular model $\boldsymbol{\gamma}_S$, we denote the p-dimensional coefficient vector by $\boldsymbol{\lambda}_S$ and denote the $|S|$-dimensional subvector of the non-zero entries of $\boldsymbol{\lambda}_S$ by $\boldsymbol{\lambda}_{\boldsymbol{\gamma}_S}$, such that $\lambda_i \neq 0$ if $i \in S$.

Review of existing fine-mapping models

We briefly review several representative fine-mapping models, including JAM, fastPAINTOR and SuSiE.

JAM.[9] Building upon prior Bayesian fine-mapping methods such as CAVIARBF[13] and FINEMAP,[10] JAM assumes the popular g-prior[29] to model the effect sizes for a given model $\boldsymbol{\gamma}_S$:

$$\begin{aligned}\boldsymbol{\lambda}_{\boldsymbol{\gamma}_S}|\sigma_y^2 &\sim \text{MVN}\left(0, g\sigma_y^2\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S}^{-1}\right),\\ \sigma_y^2 &\sim \text{Inv-Gamma}\left(0.01, 0.01\right),\end{aligned}$$

where g is a pre-determined constant. JAM assigns a beta-binomial distribution as the prior distribution of the model size, i.e., the number of the assumed causal SNPs in a model:

$$\begin{aligned}\Pr\left(\{\boldsymbol{\gamma}; \mathbf{1}\cdot\boldsymbol{\gamma} = s\}\right) &\sim \text{Beta-binomial}(s + a, p - s + b),\\ \Pr\left(\boldsymbol{\gamma}\right) &= \frac{\Pr\left(\{\boldsymbol{\gamma}; \mathbf{1}\cdot\boldsymbol{\gamma} = s\}\right)}{\binom{p}{s}},\end{aligned}$$

where $\{\boldsymbol{\gamma}; \mathbf{1}\cdot\boldsymbol{\gamma} = s\}$ is the set of models having the same dimension as s, and all models of the same size are equally likely. $a = 1$ and $b = 9$ are used in applications, corresponding to a mean proportion of truly casual SNPs of 10%.

Relative to previous fine-mapping models, JAM does not restrict the number of causal SNPs in the region and provides a computationally efficient method based on Reversible Jump MCMC to explore a wide range of candidate models. Also, unlike previous fine-mapping methods which assume that σ_y^2 is a plug-in estimate, JAM marginalizes out both $\boldsymbol{\lambda}$ and σ_y^2 to improve the power. One potential issue is that the correlation matrix $\boldsymbol{\Sigma}$ is not necessarily positive definite, and JAM cannot work properly in such cases.

fastPAINTOR[6]. fastPAINTOR uses as input Z-scores for individual variants, $Z_i = \frac{\hat{\beta}_i}{\text{se}\left(\hat{\beta}_i\right)}$, and the LD matrix $\boldsymbol{\Sigma}$, usually estimated from a reference panel. fastPAINTOR assumes the following sampling distribution for $\boldsymbol{Z}$:

$$\boldsymbol{Z}|\boldsymbol{\lambda}_S, \boldsymbol{\Sigma} \sim \text{MVN}(\boldsymbol{\Sigma}\boldsymbol{\lambda}_S, \boldsymbol{\Sigma}),$$

for a given model $\boldsymbol{\gamma}$. The assumed prior distribution of $\boldsymbol{\lambda}_S$ can be written as

$$\begin{aligned}\boldsymbol{\lambda}_S|\boldsymbol{\gamma}, \sigma_{\boldsymbol{\lambda}}^2 &\sim \text{MVN}(0, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}}),\\ \boldsymbol{\Sigma}_{\boldsymbol{\gamma}} &= \sigma_{\boldsymbol{\lambda}}^2\text{Diag}(\boldsymbol{\gamma}) + \text{Diag}(\sigma_y^2).\end{aligned}$$

The distribution of $\boldsymbol{Z}$ after integrating out $\boldsymbol{\lambda}_S$ is

$$\boldsymbol{Z}|\boldsymbol{\gamma}, \boldsymbol{\Sigma} \sim \text{MVN}(0, \boldsymbol{\Sigma} + \boldsymbol{\Sigma}\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}\boldsymbol{\Sigma})P(\boldsymbol{\gamma}).$$

The prior distribution of γ_i is a Bernoulli distribution, where the probability of being causal can be related to functional annotations as follows:

$$\gamma_i \sim \text{Bern}(\frac{\exp\{\boldsymbol{w}_i'\boldsymbol{\theta}\}}{1+\exp\{\boldsymbol{w}_i'\boldsymbol{\theta}\}}),$$

where $\boldsymbol{w}_i$ is the vector of functional annotations for SNP i. Posterior probabilities for each SNP are calculated using an importance sampling algorithm. Note that fastPAINTOR cannot handle large number of functional annotations, and a pre-selection of a small number of annotations may be necessary.[12]

SuSiE.[5] The SuSiE model works particularly well in situations where variables are highly correlated and the effects are sparse. It is conceptually different from the other fine-mapping models. The inference in the SuSiE model is based on multiple basic models, the so-called Single-Effect Regression (SER) models. Under the assumption of no prior information, the SER model is defined as:

$$\begin{aligned}
\boldsymbol{y} &= \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon} \sim \text{MVN}(\boldsymbol{0}, \sigma_y^2 I_n) \\
\boldsymbol{\beta} &= b\boldsymbol{\gamma} \\
b &\sim N(0, \sigma_0^2) \\
\boldsymbol{\gamma} &\sim \text{Mult}(1, \boldsymbol{\pi}),
\end{aligned}$$

where $\text{Mult}(m, \boldsymbol{\pi})$ denotes the multinomial distribution on class counts that is obtained when m samples are drawn with class probability $\boldsymbol{\pi}$. For the SER model, there is only one non-zero effect, and hence the indicator vector $\boldsymbol{\gamma}$ has only one non-zero element. SuSiE assumes by default that $\boldsymbol{\pi} = (1/p, \ldots, 1/p)$, but one could define $\boldsymbol{\pi}$ based on prior probabilities derived from functional annotations. Calculating the posterior inclusion probabilitles (PIP) $\Pr\left(\gamma_i = 1|\boldsymbol{y}, \boldsymbol{X}, \sigma_y^2, \sigma_0^2\right)$ involves fitting p univariate regression of $\boldsymbol{y}$ on the columns of $\boldsymbol{x}_i$ of $\boldsymbol{X}$.

The SER model assumes only one causal SNP. For extensions to multiple causal variants within a locus, the final SuSiE model is built based on L SER models. The idea is to introduce multiple single-effect vectors $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_L$ and construct the overall effect vector $\boldsymbol{\beta}$ as the sum of these single effects. The model is as follows:

$$\begin{aligned}
\boldsymbol{y} &= \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon} \sim \text{MVN}(\boldsymbol{0}, \sigma_y^2 I_n) \\
\boldsymbol{\beta} &= \sum_{l=1}^{L} \boldsymbol{\beta}_l \\
\boldsymbol{\beta}_l &= \boldsymbol{\gamma}_l b_l \\
b_l &\sim N(0, \sigma_0^2) \\
\boldsymbol{\gamma}_l &\sim \text{Mult}(1, \boldsymbol{\pi}).
\end{aligned}$$

SuSiE performs an iterative Bayesian stepwise selection (IBSS) algorithm to fit this model; at each iteration it uses the SER model to estimate $\boldsymbol{\beta}_l$ given current estimate of $\boldsymbol{\beta}_{l'}$ for $l' \neq l$. Specifically, it fits the SER model for $\boldsymbol{\beta}_l$ using the residual $\bar{r} \leftarrow \boldsymbol{y} - \boldsymbol{X}\sum_{l' \neq l}\boldsymbol{\beta}_{l'}$. The result of the SuSiE model consists of L fitted $\hat{\boldsymbol{\beta}}_l$, and L corresponding PIP vectors $\boldsymbol{\alpha}_l = \left\{\Pr\left(\boldsymbol{\beta}_{l,1} \neq 0|\boldsymbol{X}, \boldsymbol{y}\right), \ldots, \Pr\left(\boldsymbol{\beta}_{l,p} \neq 0|\boldsymbol{X}, \boldsymbol{y}\right)\right\}'$. Then the final PIP is defined as

$$\Pr\left(\hat{\beta}_j \neq 0|\boldsymbol{X}, \boldsymbol{y}\right) \approx 1 - \prod_{l \in \{1,\ldots,L\}} (1 - \alpha_{l,j}),$$

assuming that the $\beta_{l,j}$ are independent across $l = 1 \ldots L$. SuSiE naturally produces credible sets; a level ρ credible set is defined as a subset of variables that has probability ρ or more to contain at least one causal variable.

SuSiE+PolyFun.[12] To fully utilize the potential of the fine-mapping methods that can only take as input univariate prior causal probability, Weissbrod et al.[12] proposed to estimate prior probabilities through the S-LDSC model,[30] which essentially estimates global, genome-wide functional genomic enrichments using only summary data. Specifically, PolyFun estimates the prior probability for each SNP in proportion to per-SNP heritability estimates: $\Pr(\beta_i \neq 0|\boldsymbol{w}_i) \propto \hat{\text{var}}[\beta_i|\boldsymbol{w}_i] = \hat{\boldsymbol{\theta}}'\boldsymbol{w}_i$ where

$$\hat{\boldsymbol{\theta}} := \underset{\boldsymbol{\theta}\in R^q}{\text{argmin}} \sum_i \left[\chi_i^2 - n\sum_{j=1}^{q}\theta_j l(i,j) - nb - 1\right]^2 + \alpha||\boldsymbol{\theta}||^2;$$

q is the number of functional annotations, $\chi_i^2 = \frac{(\boldsymbol{x}_i'\boldsymbol{y})^2}{n}$ is the χ^2 statistic of SNP i, $l(i,j) = \sum_{k=1}^{p}\text{cor}(\boldsymbol{x}_i, \boldsymbol{x}_k)\, w_{j,k}$ is the LD-score for SNP i weighted by functional annotation j, and b measures the contribution of confounding biases.

In applications, SuSiE+PolyFun and FINEMAP+PolyFun were shown to lead to more causal variant discoveries relative to SuSiE without functional annotations. The advantage of PolyFun is that the contribution of a functional annotation to the heritability is estimated using genome-wide SNPs.

We summarize the main features and assumptions of commonly used fine-mapping methods in the literature in Table 3.

Proposed Bayesian fine-mapping model: CARMA

We introduce here the details of our model; we refer to it as CARMA (CAusal Robust Mapping method with Annotations). Given any model $\boldsymbol{\gamma}_S$, we assume that the summary statistics $\boldsymbol{Z}$ follow a multivariate normal distribution:

$$\boldsymbol{Z}|\boldsymbol{\lambda}_S, \boldsymbol{\Sigma} \sim \text{MVN}(\boldsymbol{\Sigma}\boldsymbol{\lambda}_S, \boldsymbol{\Sigma}).$$

We want to identify the true model that generated the summary statistics through posterior inference within a Bayesian paradigm.

Spike-and-slab prior distribution on the scaled effect sizes. We assume a spike-and-slab prior for the prior distribution of λ_i, i.e.,

$$\begin{aligned}\lambda_i &\sim \text{Normal}(0, \tau^{-1}) \text{ if } \gamma_i = 1,\\ \lambda_i &= 0 \text{ if } \gamma_i = 0.\end{aligned}$$

Specifically, given an index set S, the prior distribution of the assumed non-zero effect sizes $\boldsymbol{\lambda}_{\boldsymbol{\gamma}_S}$ corresponding to $\boldsymbol{\gamma}_S$ is

$$\boldsymbol{\lambda}_{\boldsymbol{\gamma}_S}|\tau, \boldsymbol{\gamma}_S \sim \text{MVN}(0, \frac{1}{\tau}I_{\boldsymbol{\gamma}_S}),$$

where $I_{\boldsymbol{\gamma}_S}$ is a $|S| \times |S|$ identity matrix. We may tune the value of the precision parameter τ according to empirical evidence (i.e. such that with 95% probability, a causal SNP explains less than 1% of the trait variation).

Alternatively, we can assign a prior distribution $f(\tau)$ on the precision parameter τ, resulting in more flexibility and robustness to situations where different loci may have different effect sizes and therefore explain different proportions of trait variance. By assigning a prior distribution on the mixing parameter τ, we intrinsically assign a mixture of normal distribution on the scaled effect size $\boldsymbol{\lambda}_{\boldsymbol{\gamma}_S}$. The choice of $f(\tau)$ directly impacts the computation of posterior probabilities for candidate models. Here, for several reasons including computational efficiency, we explore one specific form of $f(\tau)$, the Zellner-Siow's Cauchy prior:[31] $\tau \sim \text{Gamma}(\frac{1}{2}, \frac{1}{2})$. We focus our derivations below on the spike-and-slab prior, and include the corresponding results for the Cauchy prior in the Supplemental Material.

Prior distribution on model space. A fine-mapping locus typically contains a large number of variants which can be a challenge for model selection. Without an appropriate prior distribution, all candidate models including the finite-dimensional true model would receive very small posterior probabilities, as also noted in.[7] A penalization of the size of the selected model, through the prior distribution, is required for achieving model selection consistency in such high-dimensional regime as shown in.[32,33]

We introduce a prior distribution on model space to control the total number of causal SNPs that any candidate model assumes, which is analogous to controlling the false discovery rate. Let $|S| = \sum_{\gamma_i \in \boldsymbol{\gamma}_S} \gamma_i$ denote the total number of causal SNPs for a given $\boldsymbol{\gamma}_S$. We first place a discrete prior distribution on the random variable $|S|$ and let $\Pr(|S||\eta)$ denote the p.m.f. with η as a hyperparameter. Given $|S|$, we assume that all those models have the same prior probability. Hence, the prior probability of $\boldsymbol{\gamma}_S$ is

$$\Pr(\boldsymbol{\gamma}_S|\eta) = \frac{\Pr(|S||\eta)}{\binom{p}{|S|}}.$$

We propose to use the truncated Poisson prior ($|S| \in \{0, \ldots, p\}$), which has been introduced before and shown to enjoy model selection consistency in.[33] Let $F(\cdot|\eta)$ be the cumulative distribution function of a Poisson distribution with mean η, the truncated Poisson prior is defined as follows:

$$\Pr(|S||\eta) = \frac{\eta^{|S|}\exp\{-\eta\}}{|S|!F(p|\eta)} \propto \frac{\eta^{|S|}\exp\{-\eta\}}{|S|!}.$$

Since the total number of SNPs p is usually a large number and η is chosen to be small in order to reflect the sparse scenario for the true causal SNPs, $F(p|\eta) \to 1$. Then given any specific model $\boldsymbol{\gamma}_S$, the prior probability of this model under Poisson distribution is

$$\Pr(\boldsymbol{\gamma}_S|\eta) \propto \frac{\eta^{|S|}\exp\{-\eta\}(p-|S|)!}{p!}.$$

The Poisson distribution on model space imposes necessary sparsity for multiplicity control. The hyperparameter η plays a critical role in this sparsity encouraging mechanism. Note that the computation of the PIPs is based on the unnormalized posterior probabilities for the candidate models, i.e., $\Pr(\boldsymbol{\gamma}_S|\boldsymbol{Z}) \propto f(\boldsymbol{Z}|\boldsymbol{\gamma}_S)\Pr(\boldsymbol{\gamma}_S|\eta)$. Given that η can be thought of as the prior expectation of the number of causal SNPs at a given locus, we recommend $\eta = 1$ as a default choice and all our analyses are performed under this assumption.

Marginal likelihood and posterior probability. Given the Z-scores $\boldsymbol{Z}$, LD correlation matrix $\boldsymbol{\Sigma}$, and a non-empty index set S, the marginal likelihood conditional on $\boldsymbol{\gamma}_S$ is

$$m(\boldsymbol{Z}|\boldsymbol{\gamma}_S, \boldsymbol{\Sigma}) = \int_{\boldsymbol{\lambda}_S} f(\boldsymbol{Z}|\boldsymbol{\lambda}_S, \boldsymbol{\Sigma}) f(\boldsymbol{\lambda}_S|\boldsymbol{\gamma}_S, \tau) \mathrm{d}\boldsymbol{\lambda}_S,$$

where $f(\boldsymbol{Z}|\boldsymbol{\lambda}_S, \boldsymbol{\Sigma})$ is the density function of $\mathrm{MVN}(\boldsymbol{\Sigma}\boldsymbol{\lambda}_S, \boldsymbol{\Sigma})$, and $f(\boldsymbol{\lambda}_S|\boldsymbol{\gamma}_S, \tau)$ is the prior density function of the scaled effect size, which is the product of $p - |S|$ point mass distribution at 0 and the density function of $\mathrm{MVN}(0, \frac{1}{\tau} I_S)$. The marginal likelihood after integrating out $\boldsymbol{\lambda}_{\boldsymbol{\gamma}_S}$ is

$$m(\boldsymbol{Z}|\boldsymbol{\gamma}_S, \boldsymbol{\Sigma}) = |\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S} + \tau I_{\boldsymbol{\gamma}_S}|^{-\frac{1}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} (2\pi)^{-\frac{p}{2}} \tau^{\frac{|S|}{2}} \exp\left\{ -\frac{\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z} - \boldsymbol{Z}_S'(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S} + \tau I_{\boldsymbol{\gamma}_S})^{-1}\boldsymbol{Z}_S}{2} \right\}, \tag{1}$$

and the ratio between the marginal likelihood of $\boldsymbol{\gamma}_S$ and the null model $\boldsymbol{\gamma}_0$ is

$$\frac{m(\boldsymbol{Z}|\boldsymbol{\gamma}_S)}{m(\boldsymbol{Z}|\boldsymbol{\gamma}_0)} = |\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S} + \tau I_{\boldsymbol{\gamma}_S}|^{-\frac{1}{2}} \tau^{\frac{|S|}{2}} \exp\left\{ \frac{\boldsymbol{Z}_S'(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S} + \tau I_{\boldsymbol{\gamma}_S})^{-1}\boldsymbol{Z}_S}{2} \right\}.$$

The details can be found in the Supplemental Material. Let $\mathcal{M}$ denote the model set that contains all candidate models. Then the posterior probability of any non-null model $\boldsymbol{\gamma}_S$ and the posterior probability of γ_i being equal to 1 (the PIP) can be computed as

$$\Pr(\boldsymbol{\gamma}_S|\boldsymbol{Z}) = \frac{PO_{\boldsymbol{\gamma}_S:\boldsymbol{\gamma}_0}}{\sum_{\boldsymbol{\gamma}_A\in\mathcal{M}} PO_{\boldsymbol{\gamma}_A:\boldsymbol{\gamma}_0}},$$
$$\Pr(\gamma_i = 1|\boldsymbol{Z}) = \sum_{\boldsymbol{\gamma}_S:i\in S} \Pr(\boldsymbol{\gamma}_S|\boldsymbol{Z}),$$

where the posterior odds ($PO_{\boldsymbol{\gamma}_S:\boldsymbol{\gamma}_0}$) is defined as the product of the Bayes factor $\left(\frac{f(\boldsymbol{Z}|\boldsymbol{\gamma}_S)}{f(\boldsymbol{Z}|\boldsymbol{\gamma}_0)}\right)$ and the prior odds $\left(\frac{\Pr(\boldsymbol{\gamma}_S|\eta)}{\Pr(\boldsymbol{\gamma}_0|\eta)}\right)$:

$$\begin{aligned} PO_{\boldsymbol{\gamma}_S:\boldsymbol{\gamma}_0} &= \frac{m(\boldsymbol{Z}|\boldsymbol{\gamma}_S)}{m(\boldsymbol{Z}|\boldsymbol{\gamma}_0)} \cdot \frac{\Pr(\boldsymbol{\gamma}_S|\eta)}{\Pr(\boldsymbol{\gamma}_0|\eta)} \\ &= \frac{\eta^{|S|}(p-|S|)!}{p!} |\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S} + \tau I_{\boldsymbol{\gamma}_S}|^{-\frac{1}{2}} \tau^{\frac{|S|}{2}} \exp\left\{\frac{\boldsymbol{Z}_S'(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S} + \tau I_{\boldsymbol{\gamma}_S})^{-1}\boldsymbol{Z}_S}{2}\right\}. \end{aligned}$$

Shotgun stochastic search algorithm. Most fine mapping methods adopt MCMC-based algorithms to explore the posterior model space. However such algorithms can be ineffective in high dimension situations because of slow convergence and inefficient proposal distribution.[16] Furthermore, due to high correlations among SNPs at a locus, MCMC algorithms may not visit perfectly correlated SNPs evenly, which leads to unequal PIPs for such SNPs. Therefore, we adopt here the Shotgun stochastic search (Shotgun) algorithm[16] for exploring the posterior distribution over model space, which has been used before in the context of fine-mapping.[10] The main benefit of using the Shotgun algorithm is that it not only records the visited candidate models, but also the neighborhood of the visited candidate models, hence the Shotgun algorithm has a more complete exploration of the areas with high marginal likelihoods in the posterior model space. This semi-exhaustive search feature alleviates the problem of unequal PIPs for perfectly correlated SNPs. This feature will also play an important role when integrating functional annotations into the computation of the proposed model. We note that running the Shotgun algorithm in this setting is feasible due to the strong dimensional penalization introduced by the prior Poisson distribution that constrains the search in the posterior model space.

Note that for a specific model denoted by an index set S, the unnormalized posterior distribution of the indicator vector $\boldsymbol{\gamma}_S$ is proportional to a product of the marginal likelihood and the prior distribution:

$$\Pr(\boldsymbol{\gamma}_S|\boldsymbol{Z}) = \frac{f(\boldsymbol{Z}|\boldsymbol{\gamma}_S)f(\boldsymbol{\gamma}_S)}{f(\boldsymbol{Z})} \propto f(\boldsymbol{Z}|\boldsymbol{\gamma}_S)f(\boldsymbol{\gamma}_S).$$

The Shotgun algorithm is an iterative procedure that exhaustively examines the following three neighborhood sets of the current model:

$$\begin{aligned} &\Gamma_-(S) := \{A : A \subset S, |S| - |A| = 1\} \text{ (one less SNP than } S), \\ &\Gamma_+(S) := \{A : A \supset S, |A| - |S| = 1\} \text{ (one more SNP than } S), \\ &\Gamma_\Leftrightarrow(S) := \{A : |S| - |A\cap S| = 1, |A| = |S|\} \text{ (models that replace one SNP in } S). \end{aligned}$$

Then, all the unnormalized posterior probabilities of the neighborhood models, i.e., $\{\Gamma_-(S)\cup\Gamma_+(S)\cup\Gamma_\Leftrightarrow(S)\}$, will be computed. To update the current model, the algorithm first randomly selects one candidate model from each neighborhood set according to the unnormalized posterior probabilities, then randomly selects the next current model from the three selected models according to the corresponding posterior probabilities. By doing so, the algorithm stochastically moves towards the high posterior area in the model space. The Shotgun algorithm is stopped when the sum of the absolute difference of the PIPs between iterations is smaller than a pre-determined threshold, and we keep the top B visited candidate models with the largest unnormalized posterior probabilities.

Remarks

The inversion of Σ. Due to high correlations among SNPs at a locus, the LD matrix is not full rank most of the time. We use the Moore–Penrose inverse of the LD matrix instead.

Posterior probability of candidate models. Most of the fine-mapping methods that have been proposed focus on the goal of variable selection, accounting for LD. One advantage of our proposed model is that in addition to PIPs, we provide the posterior probability of candidate models which can be informative about the putative causal variants at the test locus. Note that FINEMAP also provides similar posterior probabilities of candidate models.

Specifically, CARMA returns a vast library of visited candidate models that can be ranked according to the corresponding marginal likelihoods. Let $\boldsymbol{\gamma}_{(b)}$, $b = 1, \ldots, B$, denote the ranked candidate models, such as $\boldsymbol{\gamma}_{(1)}$ receives the largest marginal likelihood. We use $\boldsymbol{\gamma}_{(1)}$ as the reference model to select all other candidate models that are not significantly different from $\boldsymbol{\gamma}_{(1)}$ through a Bayesian hypothesis testing procedure as follows. The posterior odds of the hypothesis test for comparing $\boldsymbol{\gamma}_{(1)}$ to all other candidate models is

$$\mathrm{PO}_{\boldsymbol{\gamma}_{(1)}:\boldsymbol{\gamma}_{(b)}} = \frac{\Pr\left(\boldsymbol{\gamma}_{(1)}|\boldsymbol{Z}, \eta\right)}{\Pr\left(\boldsymbol{\gamma}_{(b)}|\boldsymbol{Z}, \eta\right)} = \frac{\Pr\left(\boldsymbol{Z}|\boldsymbol{\gamma}_{(1)}\right)\Pr\left(\boldsymbol{\gamma}_{(1)}|\eta\right)}{\Pr\left(\boldsymbol{Z}|\boldsymbol{\gamma}_{(b)}\right)\Pr\left(\boldsymbol{\gamma}_{(b)}|\eta\right)},$$

for $b = 1, \ldots, B$. The posterior odds quantifies the strength of evidence in favor of the leading causal configuration, represented by $\boldsymbol{\gamma}_{(1)}$ relative to other candidate models. If all candidate models have the same prior, then the posterior odds is equal to the Bayes factor. We use commonly accepted thresholds for $\mathrm{BF}_{\boldsymbol{\gamma}_{(1)}:\boldsymbol{\gamma}_{(b)}}$ as the threshold for $PO_{\boldsymbol{\gamma}_{(1)}:\boldsymbol{\gamma}_{(b)}}$ to determine if a candidate model $\gamma_{(b)}$ is significantly different from the reference model $\gamma_{(1)}$.[18] We only keep those models with the corresponding Bayes factor smaller than the threshold (e.g. 3.2 or 10), and refer to the corresponding SNPs identified by these selected models as the credible model.

Incorporating functional annotations. Leveraging functional annotation data on genetic variants at a locus of interest may improve the results of fine-mapping studies.[6,12] Most fine-mapping methods allow inclusion of only one functional annotation in the form of the prior probability for each SNP. However, it is difficult to choose one best annotation. fastPAINTOR cannot deal with high-dimensional functional annotation data and a pre-selection of a small number of functional annotations needs to be performed before running fastPAINTOR. SuSiE+PolyFun is essentially a two-step procedure: first, it estimates prior probabilities using the S-LDSC regression framework, and then those priors are provided as input to SuSiE.

Here we propose to incorporate large number of functional annotations using a logistic distribution to model the relationship between the annotations and the prior probability, such as $\mathrm{logit}\left(\Pr\left(\gamma_i = 1|\boldsymbol{w}_i, \boldsymbol{\theta}\right)\right) = \boldsymbol{w}_i'\boldsymbol{\theta}$, where $\boldsymbol{w}_i$ is the $q+1$ dimensional vector (for q annotations) of the ith SNP. Since the indicator vector $\boldsymbol{\gamma}$ is unknown to us, an EM algorithm is typically used to replace the unknown indicator by its expectation conditional on summary statistics, functional annotations, and the estimated coefficients during the computation. Here we propose an EM algorithm with regularization on the coefficients to integrate information on large number of annotations. By adding the functional annotations, the log-likelihood can be written as:

$$\begin{aligned}\ell(\boldsymbol{\theta}; \boldsymbol{Z}, \boldsymbol{\gamma}, W) &= \log\left(m(\boldsymbol{Z}|\boldsymbol{\gamma})\right) + \log\left(\Pr\left(\boldsymbol{\gamma}|\boldsymbol{W}, \boldsymbol{\theta}\right)\right) \\ &= \log\left(m(\boldsymbol{Z}|\boldsymbol{\gamma})\right) + \sum_{i=1}^{p}\left[\gamma_i \boldsymbol{w}_i'\boldsymbol{\theta} - (1-\gamma_i)\log\left(1+\exp(\boldsymbol{w}_i'\boldsymbol{\theta})\right)\right], \end{aligned} \tag{2}$$

where $m(\boldsymbol{Z}|\boldsymbol{\gamma})$ is given by Equation (1). The EM algorithm is as follows:

Input: Summary statistics $\boldsymbol{Z}$, functional annotations $\boldsymbol{W}$, hyperparameter η of the Poisson prior distribution (by default $\eta = 1$).
Initialization: Run Shotgun algorithm without annotations to generate the initial posterior probability $\Pr(\boldsymbol{\gamma}|\boldsymbol{Z}, \eta)$.
for $s = 1, \ldots$ until convergence **do**

E-step
- Run Shotgun algorithm with initial prior probability $\Pr\left(\gamma_i|\boldsymbol{w}_i, \boldsymbol{\theta}^{(s)}\right) = \text{logit}^{-1}\left(\boldsymbol{w}_i'\boldsymbol{\theta}^{(s)}\right)$ estimated at step $(s-1)$.
- Compute the posterior probability $\Pr\left(\gamma_i = 1|\boldsymbol{Z}, \boldsymbol{W}, \boldsymbol{\theta}^{(s)}\right)$:

$$\Pr\left(\boldsymbol{\gamma}_S|\boldsymbol{Z}, \boldsymbol{W}, \boldsymbol{\theta}^{(s)}\right) = \frac{m(\boldsymbol{Z}|\boldsymbol{\gamma}_S)\Pr\left(\boldsymbol{\gamma}_S|\boldsymbol{W}, \boldsymbol{\theta}^{(s)}\right)}{\sum_{\boldsymbol{\gamma}_A \in \mathcal{M}} m(\boldsymbol{Z}|\boldsymbol{\gamma}_A)\Pr\left(\boldsymbol{\gamma}_A|\boldsymbol{W}, \boldsymbol{\theta}^{(s)}\right)},$$
$$\Pr\left(\gamma_i = 1|\boldsymbol{Z}, \boldsymbol{W}, \boldsymbol{\theta}^{(s)}\right) = \sum_{\boldsymbol{\gamma}_S : i \in S} \Pr\left(\boldsymbol{\gamma}_S|\boldsymbol{Z}, \boldsymbol{W}, \boldsymbol{\theta}^{(s)}\right).$$

- Compute the Q function, i.e. the expectation of the log-likelihood with respect to the posterior probability of $\boldsymbol{\gamma}$ as follows:

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\gamma}, \boldsymbol{\theta}^{(s)}) &= \mathbf{E}_{\boldsymbol{\gamma}|\boldsymbol{Z}, \boldsymbol{W}, \boldsymbol{\theta}^{(s)}}\left[\ell(\boldsymbol{\theta}; \boldsymbol{Z}, \boldsymbol{\gamma}, \boldsymbol{W})\right] \\ &= \mathbf{E}_{\boldsymbol{\gamma}|\boldsymbol{Z}, \boldsymbol{W}, \boldsymbol{\theta}^{(s)}}\left[\log\left(m(\boldsymbol{Z}|\boldsymbol{\gamma})\right)\right] + \sum_{i=1}^{p}\left[p_i(\boldsymbol{\theta}^{(s)})\boldsymbol{w}_i'\boldsymbol{\theta} - \left(1 - p_i(\boldsymbol{\theta}^{(s)})\right)\log\left(1 + \exp(\boldsymbol{w}_i'\boldsymbol{\theta})\right)\right], \quad (3) \end{aligned}$$

where we denote by

$$p_i(\boldsymbol{\theta}^{(s)}) = \Pr\left(\gamma_i = 1|\boldsymbol{Z}, \boldsymbol{W}, \boldsymbol{\theta}^{(s)}\right).$$

M-step We wish to estimate $\boldsymbol{\theta}$ by maximizing the Q function with regularization on $\boldsymbol{\theta}$ (except for the intercept) to prevent overfitting with large number of annotations. We use cyclical coordinate descent algorithm[34] for finding the maximizer. We maximize the penalized Q function as follows:

$$\boldsymbol{\theta}^{(s+1)} := \underset{\boldsymbol{\theta} \in R^{q+1}}{\text{argmax}} \sum_{i=1}^{p}\left[p_i(\boldsymbol{\theta}^{(s)})\boldsymbol{w}_i'\boldsymbol{\theta} - \left(1 - p_i(\boldsymbol{\theta}^{(s)})\right)\log\left(1 + \exp\{\boldsymbol{w}_i'\boldsymbol{\theta}\}\right)\right] - \kappa P_a(\boldsymbol{\theta}_{-0}),$$

where

$$P_a(\boldsymbol{\theta}_{-0}) = \frac{(1-a)}{2}||\boldsymbol{\theta}_{-0}||_{\ell_2}^2 + a||\boldsymbol{\theta}_{-0}||_{\ell_1},$$

and $\boldsymbol{\theta}_{-0}$ is the coefficient vector without intercept, i.e. the coefficient vector of the q annotations. The tuning parameter κ is chosen by a standard cross-validation procedure based on the deviance. Notice that in this step we can drop the first term in Equation (3) because there is no $\boldsymbol{\theta}$ involved in the maximization. We assume $a = 0$ in our implementation (i.e. ridge penalty).

Then, compute the estimated prior probability for the $(s+1)$ step:

$$\Pr\left(\gamma_i = 1|\boldsymbol{w}_i, \boldsymbol{\theta}^{(s+1)}\right) = \frac{\exp(\boldsymbol{w}_i'\boldsymbol{\theta}^{(s+1)})}{1 + \exp(\boldsymbol{w}_i'\boldsymbol{\theta}^{(s+1)})},$$

for $i = 1, \ldots, p$.

end

Algorithm 1: EM algorithm with functional annotations.

Remarks

Feature selection on functional annotations. As in PolyFun, we use the ridge penalty ($a = 0$) to regularize the strength of the coefficients of annotations. If the dimension of annotations is very high, an alternative penalty is to use the elastic net penalty ($a = 0.5$).

Identifying SNPs with Z-score/LD discrepancies. A challenge when doing fine-mapping with summary statistics and LD from external reference panels is the possibility of discrepancies between LD and Z-score values. This can happen for several reasons. For example, when using a reference panel to estimate the LD, or when summary statistics at SNPs in a region are based on studies with different sample sizes. Such scenarios can result in unrealistically large values for the marginal likelihood and high PIP values, especially given the large sample sizes for meta-analysis studies. Here, we propose a Bayesian hypothesis testing procedure for outlier detection embedded in the Shotgun algorithm.

Bayesian hypothesis testing. Suppose that M_S is the currently selected model in the Shotgun algorithm with a non-null index set S corresponding to $\boldsymbol{\gamma}_S$. Let $\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S}$ denote the corresponding LD matrix. When a discrepancy is identified in the index set S, the quadratic term $\boldsymbol{Z}_S'\big(\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S} + \tau I_{\boldsymbol{\gamma}_S}\big)^{-1}\boldsymbol{Z}_S$ will be substantially inflated, which biases the corresponding marginal likelihood $m(\boldsymbol{Z}|\boldsymbol{\gamma}_S, \boldsymbol{\Sigma})$ and leads to biased posterior analysis.

The proposed outlier detection procedure is based on the work in,[35] where the authors proposed a ridge regularization of the correlation matrix that leads to a stable likelihood-based test. Specifically, for a given LD matrix $\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S}$, we construct another LD matrix $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\gamma}_S}$ such as

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\gamma}_S} = \xi\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S} + (1-\xi)\mathbf{I}, \text{ for } \xi \in [0,1],$$

where $\mathbf{I}$ is the identity matrix of same dimension as $\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S}$ and for which we can compute $m(\boldsymbol{Z}_S|\boldsymbol{\gamma}_S, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\gamma}_S})$. We conduct hypothesis testing as follows:

$$H_0 : \xi = 1 (\text{there is no discrepancy}),$$
$$H_1 : \xi \neq 1 (\text{there is discrepancy}).$$

Given an equal prior probability on the null and alternative hypothesis, we derive the Bayes factor of the test as a function of ξ following the robust Bayesian approach:[36]

$$BF_{0:1} = \inf_{\xi\in[0,1]} B(\xi) = \frac{m(\boldsymbol{Z}_S|\boldsymbol{\gamma}_S, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_S})}{\sup_{\xi\in[0,1]} m(\boldsymbol{Z}_S|\boldsymbol{\gamma}_S, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\gamma}_S})},$$

where

$$\hat{\xi} := \operatorname*{argmax}_{\xi\in[0,1]} m(\boldsymbol{Z}_S|\boldsymbol{\gamma}_S, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\gamma}_S}).$$

We reject the null hypothesis, i.e. the marginal likelihood is indeed inflated by the presence of discrepancies, if $BF_{0:1} < \delta$. Possible values for the threshold δ are suggested in,[18] and we assess sensitivity to different choices. We find the maximizer $\hat{\xi}$ using the R function `optimize`. (More discussion is in the Supplemental Material.)

Imputation before fine-mapping. One way to mitigate the effect of differential sample sizes across SNPs in a meta-analysis study is to impute the Z-scores at SNPs with small sample sizes using data from SNPs with larger sample sizes. Here, we use a simple imputation method similar to the one used in.[1] Let A and B denote the index set of two cohorts with smaller and larger sample sizes, respectively; then we can impute the Z-scores of the SNPs in set A, denoted as $\boldsymbol{Z}_A$, conditional on the LD matrix $\boldsymbol{\Sigma}$ and $\boldsymbol{Z}_B$. In,[1] the authors suggest to approximate $\mathbf{E}[Z_i]$, for $i \in A$, by $\Sigma_{i,c}Z_c$, where Z_c is the leading Z-score at the locus and $\Sigma_{i,c}$ is the Pearson correlation between Z_i and Z_c. Similarly, for a given SNP $i \in A$, we propose to impute Z_i by

$$\widehat{Z}_i = \Sigma_{i,m}Z_m,$$

where $m = \operatorname*{argmax}_{j\in B} \{|\Sigma_{i,j}Z_j|\}$. Note that much more advanced imputation methods can be used instead, and our simple implementation is used here only as a proof-of-principle.

Implementation in the Shotgun algorithm. Suppose that M_S is the currently selected model in the Shotgun algorithm with a non-null index set S (note that M_S has already passed the outlier hypothesis test in the previous step of the Shotgun algorithm, therefore we assume that there are no discrepancies included in M_S for the hypothesis test at the current step). There are three sets of candidate models in the neighborhood space of M_S, i.e. $\Gamma_-(S)$, $\Gamma_+(S)$, and $\Gamma_\Leftrightarrow(S)$ as defined in the section on the Shotgun algorithm. The Shotgun algorithm selects one candidate model from each set according to the posterior probabilities, which could be inflated by the outliers. Therefore, for each selected model from the model sets $\Gamma_+(S)$ and $\Gamma_\Leftrightarrow(S)$, we conduct the Bayesian hypothesis test described above to examine whether the selection is due to inflated marginal likelihood caused by possible outliers included in the selected models. Note that the candidate models included in $\Gamma_-(S)$ have already been tested when the algorithm reached the currently selected model $\boldsymbol{\gamma}_S$, therefore we do not need to test the models in set $\Gamma_-(S)$. We drop the included/exchanged SNP from the algorithm if the corresponding Bayes factor is smaller than a threshold δ. Algorithm 2 shows the exact procedure of the outlier detection.

Note that, for a current model S, the Shotgun algorithm is a semi-exhaustive searching algorithm that considers all configurations associated with the SNPs in S. Furthermore, the Shotgun algorithm stochastically moves towards the area of large posterior probabilities in the model space, therefore the outlier detection procedure is more efficient at detecting outliers that could bias the fine-mapping analysis.

At any step of the Shotgun algorithm, suppose that the current model is $\boldsymbol{\gamma}_S$ and $S \neq \emptyset$.
Input: The index set S for the current model $\boldsymbol{\gamma}_S$, the threshold on the Bayes factor δ.

Construct two neighborhood model sets $\Gamma_+(S)$ and $\Gamma_\Leftrightarrow(S)$:

$$\Gamma_+(S) := \{A : A \supset S, |A| - |S| = 1\} \text{ (one more SNP than } S),$$
$$\Gamma_\Leftrightarrow(S) := \{A : |S| - |A \cap S| = 1, |A| = |S|\} \text{ (models that replaces one SNP in } S).$$

for $\Gamma_+(S)$ and $\Gamma_\Leftrightarrow(S)$ **do**
  repeat
    - Randomly select one candidate model A from the neighborhood model set according to the unnormalized posterior probabilities.
    - $i = A \setminus (A \bigcap S)$.
    - Define the hypothesis test for the marginal likelihood of A, such that

$$H_0 : \xi = 1; \qquad i\text{th SNP is not an outlier}$$
$$H_1 : \xi \neq 1; \qquad i\text{th SNP is an outlier.}$$

    - Compute the corresponding Bayes factor

$$\hat{B}_A = \inf_{\xi \in [0,1]} B(\xi) = \frac{m(\boldsymbol{Z}|\boldsymbol{\gamma}_A, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_A})}{\sup_{\xi \in [0,1]} m(\boldsymbol{Z}|\boldsymbol{\gamma}_A, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\gamma}_A})},$$

      where $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\gamma}_A} = \xi \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_A} + (1 - \xi)\mathbf{I}$

    if $\hat{B}_A \leq \delta$ **then**
      - Drop ith SNP from consideration for the fine-mapping algorithm.

  until $\hat{B}_A > \delta$;

end

Algorithm 2: The outlier detection procedure implemented in Shotgun algorithm.

Multi-locus extension

As shown before,[6,12,37] a fine-mapping method may benefit from inference based on multiple loci. The multiple locus model assumes that functional annotations are similarly associated with the probability to be causal. In our applications we fit the proposed model at the chromosome level.

Credible sets

In[5] the authors define a credible set as follows.

Definition 1. *A level ρ credible set is defined to be a subset of correlated variables (with correlation within the set greater than some threshold r) that has probability ρ or greater of containing at least one effect variable (i.e. causal SNP).*

Constructing credible sets. Given a correlation threshold r, we can define multiple candidate credible sets as follows

$$S := \{i \in \{1, \ldots, p\} : \min \{cor(i, j) \geq r\}, \text{ for all } i, j \in S\}.$$

Given an index set S, let $s = \left\{s_1, \ldots, s_{|S|}\right\}$ denote the indices of the selected SNPs in the set S ranked in order of decreasing PIPs, such as $\Pr(\gamma_{s_1}|\boldsymbol{Z}) > \Pr(\gamma_{s_2}|\boldsymbol{Z}) > \ldots > \Pr\left(\gamma_{s_{|S|}}|\boldsymbol{Z}\right)$, and let P_l denote the cumulative sum of the l largest PIPs:

$$P_l = \sum_{j=1}^{l} \Pr\left(\gamma_{s_j}|\boldsymbol{Z}\right).$$

Then, the credible set is defined as $S_{l_0} = \{s_1, \ldots, s_{l_0}\}$, where $l_0 = \min\{l : P_l \geq \rho\}$. This procedure makes sure that the selected credible sets have the smallest size. Then, given a specific level ρ, we run this procedure on a predetermined sequence of candidate correlation values, such as $r \in \{0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95\}$, and compute the corresponding number of credible sets for each candidate value. The largest correlation threshold r with level ρ credible sets (those sets with posterior probability greater than ρ) is selected along with the corresponding credible sets.

Data Availability. LD matrices for UK Biobank British-ancestry samples are precomputed by the PolyFun study and accessed from https://data.broadinstitute.org/alkesgroup/UKBB_LD. Baseline-LF v2.2.UKB annotations for 19 million imputed UK Biobank SNPs are available from https://data.broadinstitute.org/alkesgroup/LDSCORE/baselineLF_v2.2.UKB.tar.gz. Summary statistics of AD GWAS are available at https://ctg.cncr.nl/software/summary_statistics. The 1000 Genomes Project phase 3 genotype data are available through ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/.

Code Availability. CARMA has been implemented in a computationally efficient R package available at https://github.com/Iuliana-Ionita-Laza/CARMA.

References

[1] Masahiro Kanai, Roy Elzur, Wei Zhou, Mark J Daly, Hilary K Finucane, Global Biobank Meta analysis Initiative, et al. Meta-analysis fine-mapping is often miscalibrated at single-variant resolution. medRxiv, 2022.

[2] Christian Benner, Aki Havulinna, Marjo-Riitta Järvelin, Veikko Salomaa, Samuli Ripatti, and Matti Pirinen. Prospects of fine-mapping trait-associated genomic regions by using summary statistics from genome-wide association studies. Am J Hum Genet., (101):539–551, 2017.

[3] Iris E Jansen, Jeanne E Savage, Kyoko Watanabe, Julien Bryois, Dylan M Williams, Stacy Steinberg, Julia Sealock, Ida K Karlsson, Sara Hägg, Lavinia Athanasiu, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing alzheimer's disease risk. Nature genetics, 51(3):404–413, 2019.

[4] Clare Bycroft, Colin Freeman, Desislava Petkova, Gavin Band, Lloyd T Elliott, Kevin Sharp, Allan Motyer, Damjan Vukcevic, Olivier Delaneau, Jared O'Connell, et al. The uk biobank resource with deep phenotyping and genomic data. Nature, 562(7726):203–209, 2018.

[5] Gao Wang, Abhishek K Sarkar, Peter Carbonetto, and Matthew Stephens. 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(5):1273–1300, 2020.

[6] Gleb Kichaev, Megan Roytman, Ruth Johnson, Eleazar Eskin, Sara Lindstroem, Peter Kraft, and Bogdan Pasaniuc. Improved methods for multi-trait fine mapping of pleiotropic risk loci. Bioinformatics, 33(2):248–255, 2017.

[7] Yongtao Guan and Matthew Stephens. Bayesian variable selection regression for genome-wide association studies and other large-scale problems. The Annals of Applied Statistics, pages 1780–1815, 2011.

[8] Julian B Maller, Gilean McVean, Jake Byrnes, Damjan Vukcevic, Kimmo Palin, Zhan Su, Joanna MM Howson, Adam Auton, Simon Myers, Andrew Morris, et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. Nature genetics, 44(12):1294, 2012.

[9] Paul J Newcombe, David V Conti, and Sylvia Richardson. Jam: a scalable bayesian framework for joint analysis of marginal snp effects. Genetic epidemiology, 40(3):188–201, 2016.

[10] Christian Benner, Chris CA Spencer, Aki S Havulinna, Veikko Salomaa, Samuli Ripatti, and Matti Pirinen. Finemap: efficient variable selection using summary data from genome-wide association studies. Bioinformatics, 32(10):1493–1501, 2016.

[11] Gleb Kichaev, Wen-Yun Yang, Sara Lindstrom, Farhad Hormozdiari, Eleazar Eskin, Alkes L Price, Peter Kraft, and Bogdan Pasaniuc. Integrating functional data to prioritize causal variants in statistical fine-mapping studies. PLoS Genet, 10(10):e1004722, 2014.

[12] Omer Weissbrod, Farhad Hormozdiari, Christian Benner, Ran Cui, Jacob Ulirsch, Steven Gazal, Armin P Schoech, Bryce Van De Geijn, Yakir Reshef, Carla Márquez-Luna, et al. Functionally informed fine-mapping and polygenic localization of complex trait heritability. Nature Genetics, pages 1–9, 2020.

[13] Wenan Chen, Beth R Larrabee, Inna G Ovsyannikova, Richard B Kennedy, Iana H Haralambieva, Gregory A Poland, and Daniel J Schaid. Fine mapping causal variants with an approximate bayesian method using marginal test statistics. Genetics, 200(3):719–736, 2015.

[14] James G Scott and James O Berger. Bayes and empirical-bayes multiplicity adjustment in the variable-selection problem. The Annals of Statistics, pages 2587–2619, 2010.

[15] Jian Yang, Teresa Ferreira, Andrew P Morris, Sarah E Medland, Pamela AF Madden, Andrew C Heath, Nicholas G Martin, Grant W Montgomery, Michael N Weedon, Ruth J Loos, et al. Conditional and joint multiple-snp analysis of gwas summary statistics identifies additional variants influencing complex traits. Nature genetics, 44(4):369–375, 2012.

[16] Chris Hans, Adrian Dobra, and Mike West. Shotgun stochastic search for "large p" regression. Journal of the American Statistical Association, 102(478):507–516, 2007.

[17] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. Journal of the royal statistical society: series B (statistical methodology), 67(2):301–320, 2005.

[18] Robert E Kass and Adrian E Raftery. Bayes factors. Journal of the american statistical association, 90(430):773–795, 1995.

[19] Apostolos Dimitromanolakis, Jingxiong Xu, Agnieszka Krol, and Laurent Briollais. sim1000g: a user-friendly genetic variant simulator in r for unrelated individuals and family-based designs. BMC bioinformatics, 20(1):26, 2019.

[20] Laura Fachal, Hugues Aschard, Jonathan Beesley, Daniel R Barnes, Jamie Allen, Siddhartha Kar, Karen A Pooley, Joe Dennis, Kyriaki Michailidou, Constance Turman, et al. Fine-mapping of 150 breast cancer risk regions identifies 191 likely target genes. Nature genetics, 52(1):56–73, 2020.

[21] Jian Zhou and Olga G Troyanskaya. Predicting effects of noncoding variants with deep learning–based sequence model. Nature methods, 12(10):931–934, 2015.

[22] Yuxin Zou, Peter Carbonetto, Gao Wang, and Matthew Stephens. Fine-mapping from summary data with the "sum of single effects" model. bioRxiv, 2021.

[23] Cristen J Willer, Yun Li, and Gonalo R Abecasis. Metal: fast and efficient meta-analysis of genomewide association scans. Bioinformatics, 26(17):2190–2191, 2010.

[24] Martin Kircher, Daniela M Witten, Preti Jain, Brian J O'roak, Gregory M Cooper, and Jay Shendure. A general framework for estimating the relative pathogenicity of human genetic variants. Nature genetics, 46(3):310–315, 2014.

[25] Iuliana Ionita-Laza, Kenneth McCallum, Bin Xu, and Joseph D Buxbaum. A spectral approach integrating functional genomic annotations for coding and noncoding variants. Nature genetics, 48(2):214, 2016.

[26] GTEx Consortium et al. The gtex consortium atlas of genetic regulatory effects across human tissues. Science, 369(6509):1318–1330, 2020.

[27] Bernard Ng, Charles C White, Hans-Ulrich Klein, Solveig K Sieberts, Cristin McCabe, Ellis Patrick, Jishu Xu, Lei Yu, Chris Gaiteri, David A Bennett, et al. An xqtl map integrates the genetic architecture of the human brain's transcriptome and epigenome. Nature neuroscience, 20(10):1418–1426, 2017.

[28] Rebecca Sims, Sven J Van Der Lee, Adam C Naj, Céline Bellenguez, Nandini Badarinarayan, Johanna Jakobsdottir, Brian W Kunkle, Anne Boland, Rachel Raybould, Joshua C Bis, et al. Rare coding variants in plcg2, abi3, and trem2 implicate microglial-mediated innate immunity in alzheimer's disease. Nature genetics, 49(9):1373–1384, 2017.

[29] Arnold Zellner. On assessing prior distributions and bayesian regression analysis with g-prior distributions. Bayesian inference and decision techniques, 1986.

[30] Hilary K Finucane, Brendan Bulik-Sullivan, Alexander Gusev, Gosia Trynka, Yakir Reshef, Po-Ru Loh, Verneri Anttila, Han Xu, Chongzhi Zang, Kyle Farh, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nature genetics, 47(11):1228, 2015.

[31] Arnold Zellner and Aloysius Siow. Posterior odds ratios for selected regression hypotheses. Trabajos de estadística y de investigación operativa, 31(1):585–603, 1980.

[32] Ismaël Castillo, Aad van der Vaart, et al. Needles and straw in a haystack: Posterior concentration for possibly sparse sequences. The Annals of Statistics, 40(4):2069–2101, 2012.

[33] Andrew J Womack, Claudio Fuentes, and Daniel Taylor-Rodriguez. Model space priors for objective sparse bayesian regression. arXiv preprint arXiv:1511.04745, 2015.

[34] Jerome Friedman, Trevor Hastie, and Rob Tibshirani. Regularization paths for generalized linear models via coordinate descent. Journal of statistical software, 33(1):1, 2010.

[35] David I Warton. Penalized normal likelihood and ridge regularization of correlation and covariance matrices. Journal of the American Statistical Association, 103(481):340–349, 2008.

[36] MJ Bayarri and James O Berger. Robust Bayesian bounds for outlier detection. De Gruyter, 1992.

[37] Wenan Chen, Shannon K McDonnell, Stephen N Thibodeau, Lori S Tillmans, and Daniel J Schaid. Incorporating functional annotations for fine-mapping causal variants in a bayesian framework using summary statistics. Genetics, 204(3):933–958, 2016.

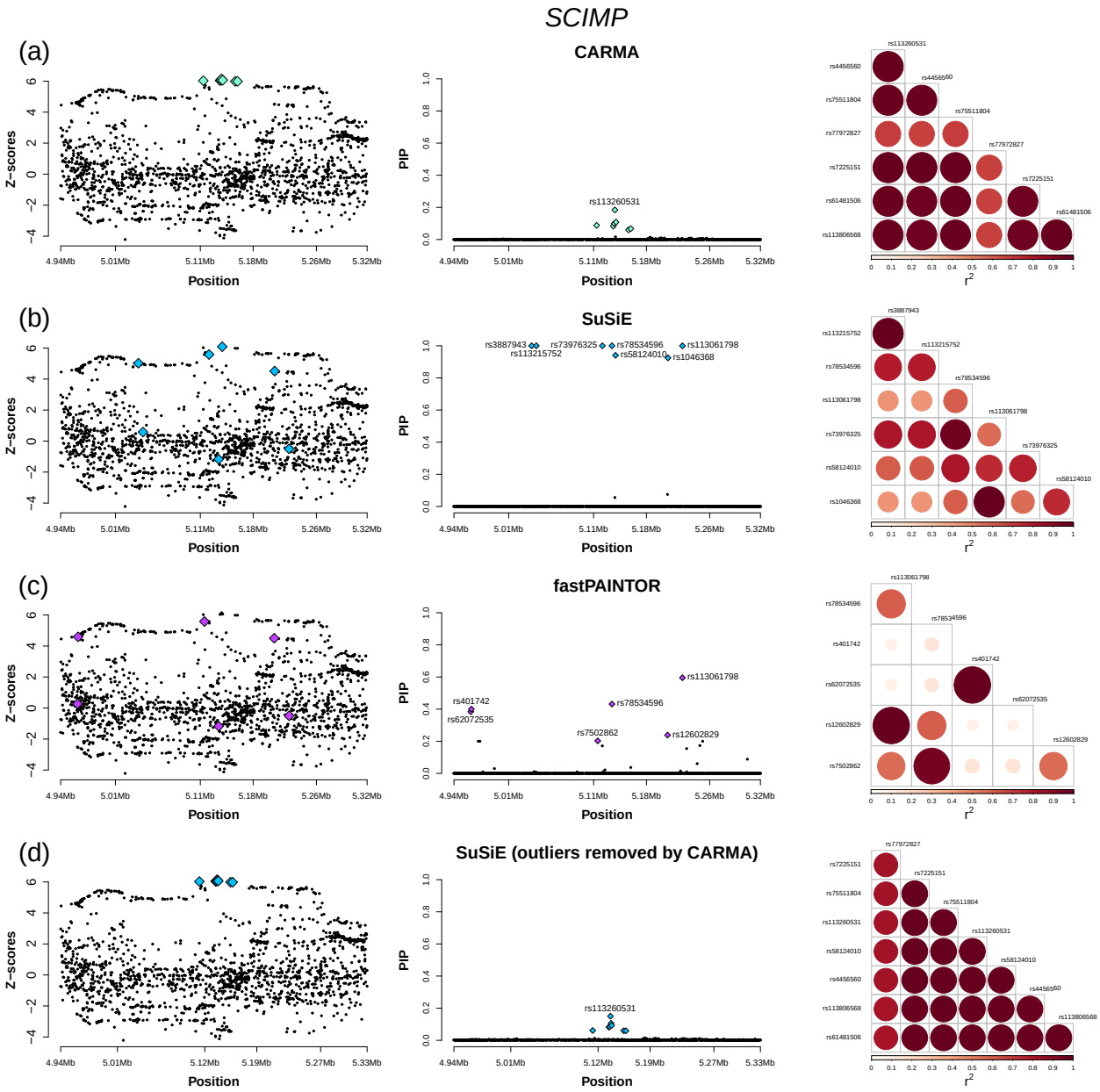

Figure 1: **Motivating example.** (a-c) GWAS Z-scores along with PIPs from three models (CARMA, SuSiE and fastPAINTOR) for one Alzheimer's disease risk locus *SCIMP* (GRCh37/hg19). For each model, the heatmap depicts the LD (r^2 based on the UKBB data) between SNPs highlighted in the PIP panel in the middle. The highlighted top SNPs of CARMA belong to one credible set; the highlighted top SNPs of SuSiE belong to seven different credible sets; the highlighted top SNPs of fastPAINTOR belong to three different credible sets. (d) Results of SuSiE applied to a dataset where outliers identified by CARMA are removed. Note that for better visualization, only a segment of the 1 Mb window analyzed at this locus is shown.

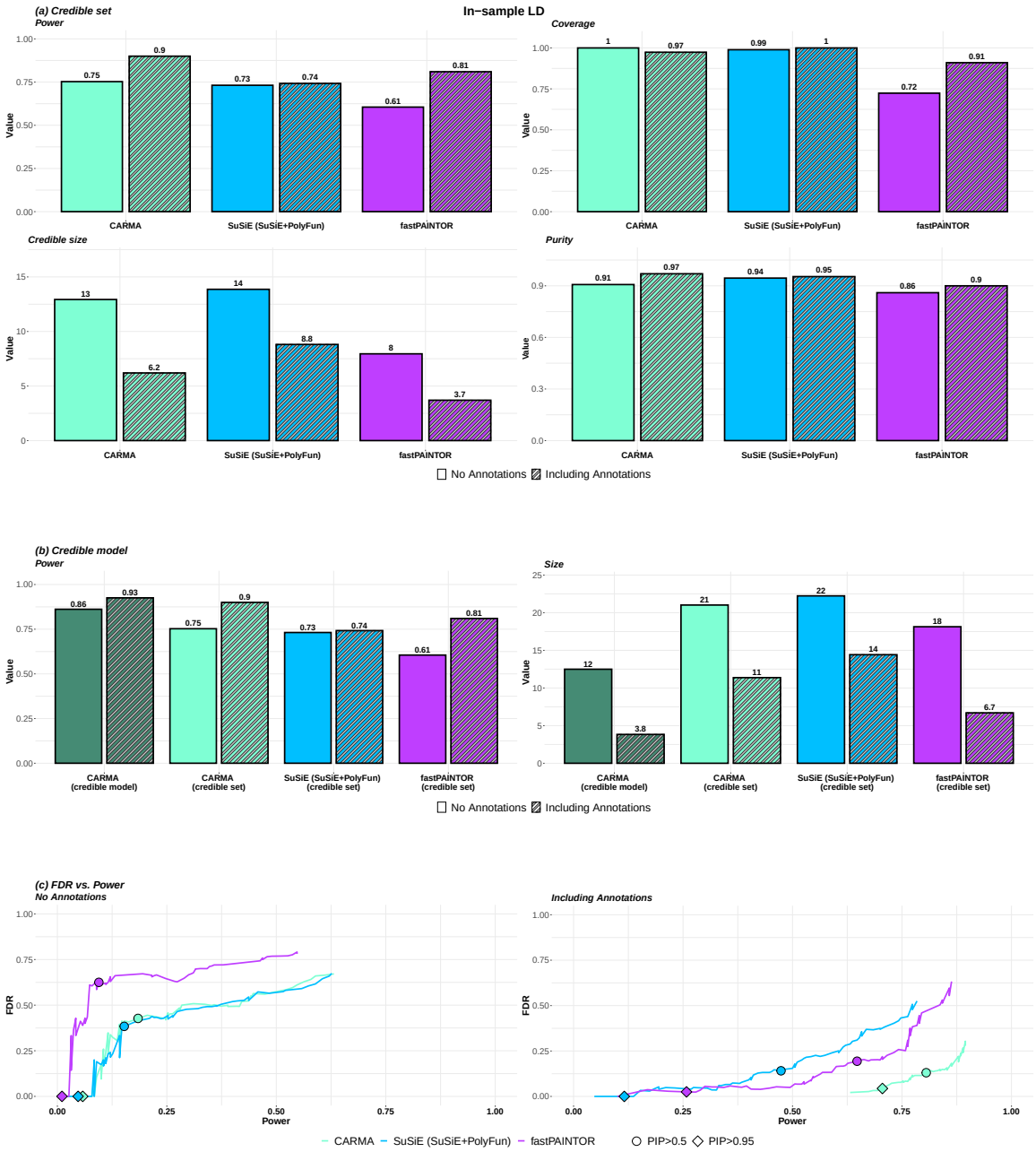

Figure 2: **Credible sets, credible models, and FDR vs. Power (in-sample LD, with and without functional annotations).** **(a)** Performance of credible sets ($\rho = 0.99$). Power: the proportion of simulated causal variants included in any credible set; Coverage: the proportion of credible sets containing a causal variant; Size: the number of variants included in a credible set; Purity: the mean squared correlation of variants in a credible set. **(b)** Performance of credible models. Power is the percentage of the simulated causal variants identified by the credible model or the credible sets. Size is the average number of variants included in the credible model (CARMA) or all the credible sets (CARMA, SuSiE, fastPAINTOR) at a locus. The credible model is computed based on a threshold of 10 for the posterior odds. **(c)** FDR vs. power using positive predictions as the PIP threshold varies from 0.1 to 1. These quantities are calculated as FDR $:= \frac{\text{FP}}{\text{TP+FP}}$ and power $:= \frac{\text{TP}}{\text{TP+FN}}$, where FP, TP, FN, TN denote the number of false positives, true positives, false negatives and true negatives respectively given a certain PIP threshold. Open circles denote the results at PIP threshold 0.5, and solid circles denote the results at PIP threshold 0.95. Two scenarios are shown: no annotation and including functional annotations. The true number of causal variants is $|T| = 2$.

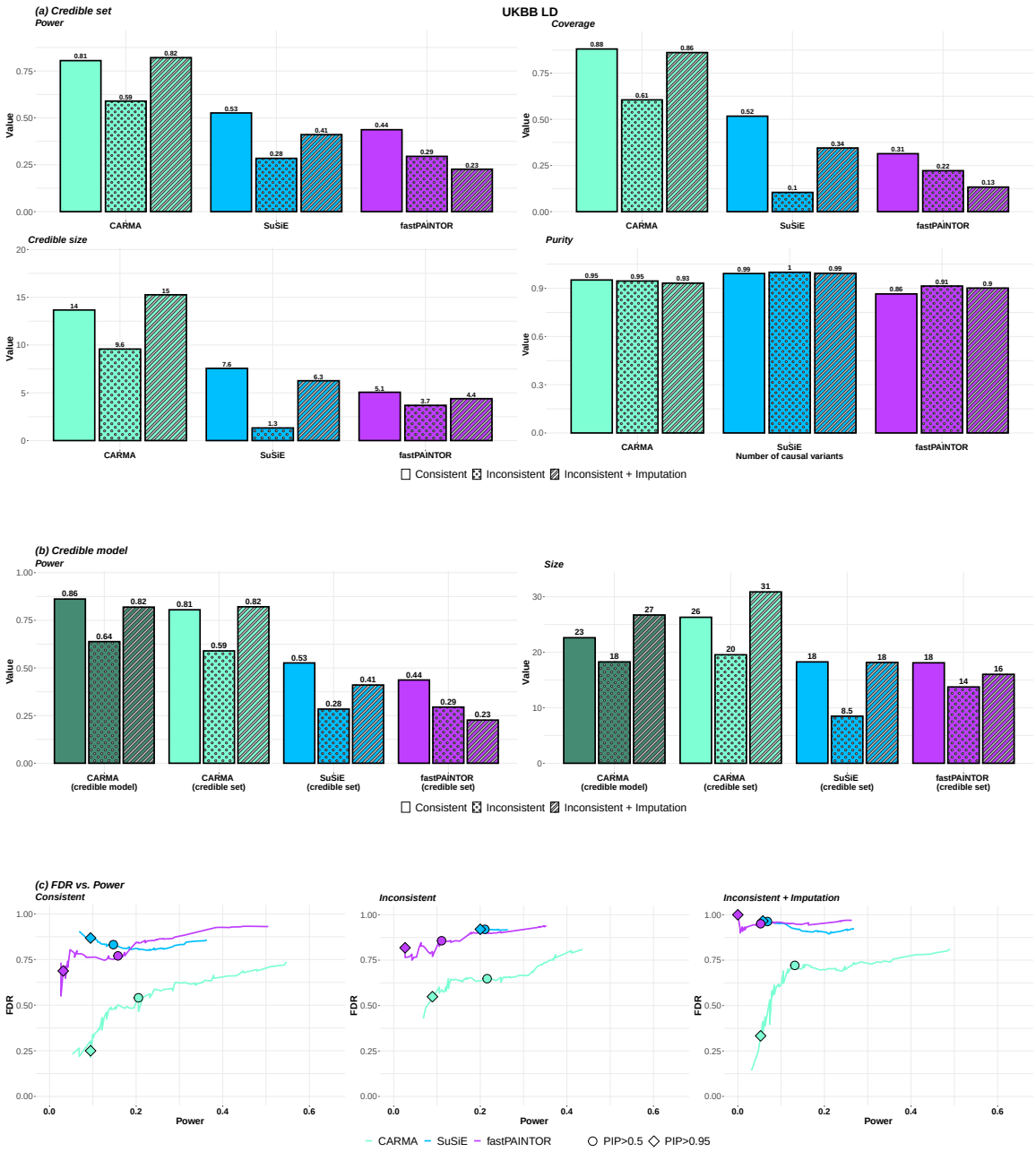

Figure 3: **Credible sets, credible models, and FDR vs. Power (UKBB LD) (a)** Performance of credible sets ($\rho = 0.99$). Power: the proportion of simulated causal variants included in any credible set; Coverage: the proportion of credible sets containing a causal variant; Size: the number of variants included in a credible set; Purity: the mean squared correlation of variants in a credible set. **(b)** Performance of credible models. Power is the percentage of the simulated causal variants identified by the credible model or the credible sets. Size is the average number of variants included in the credible model (CARMA) or all the credible sets at a locus. The credible model is computed based on a threshold of 10 for the posterior odds. **(c)** FDR vs. power using positive predictions as the PIP threshold varies from 0.1 to 1. These quantities are calculated as FDR $:= \frac{\text{FP}}{\text{TP+FP}}$ and power $:= \frac{\text{TP}}{\text{TP+FN}}$, where FP, TP, FN, TN denote the number of false positives, true positives, false negatives and true negatives respectively given a certain PIP threshold. Open circles denote the results at PIP threshold 0.5, and solid circles denote the results at PIP threshold 0.95. Two scenarios are shown: consistent and inconsistent meta-analyses. The true number of causal variants is $|T| = 2$.

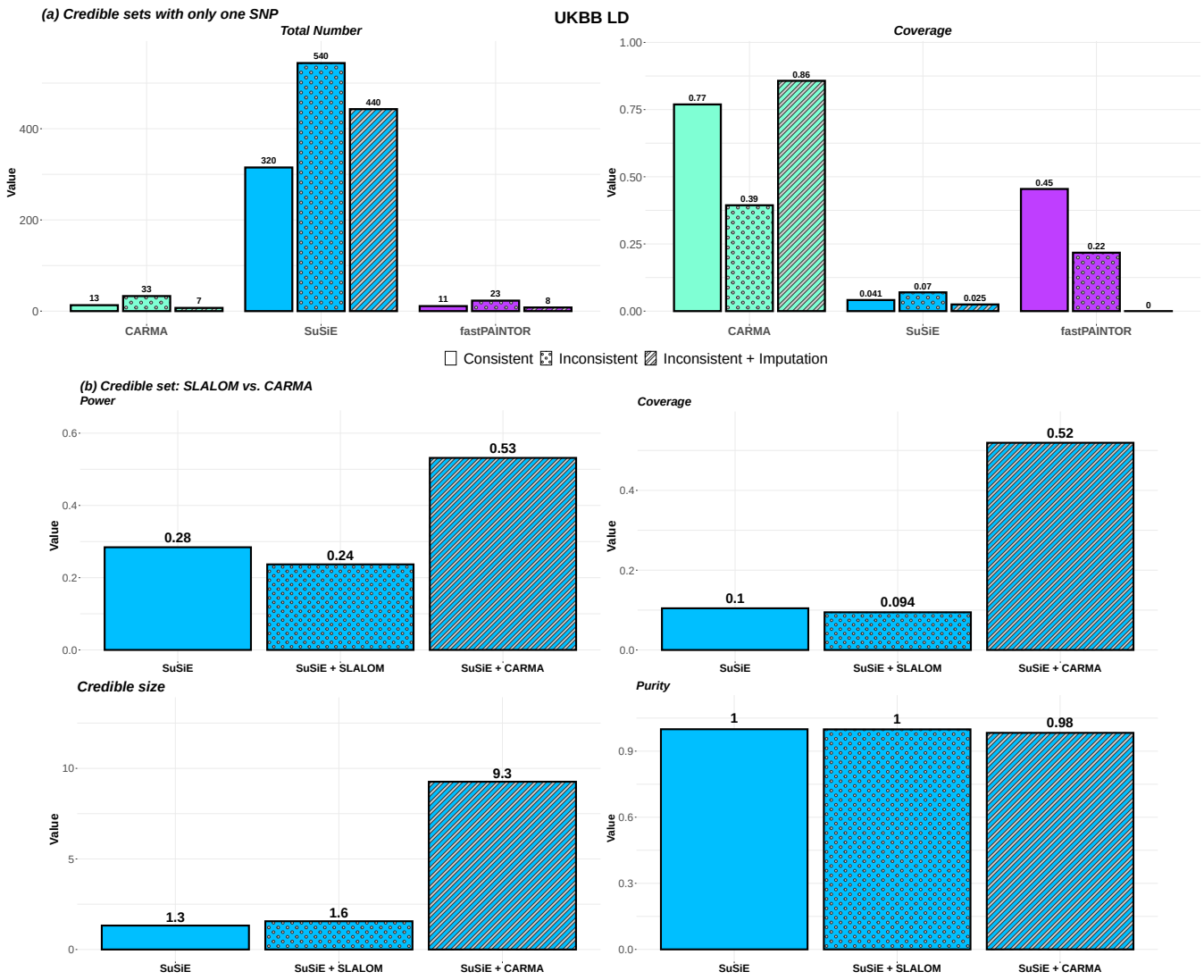

Figure 4: **(a) Increased single SNP credible sets in the presence of Z-scores/LD inconsistencies.** Results on credible sets ($\rho = 0.99$) with only one SNP, i.e., the corresponding SNP receives a PIP larger than 0.99. For each model we report the total number of credible sets that contain only one SNP across all 94 loci and three different scenarios, and the coverage of these sets, i.e. the proportion of these credible sets that contain a causal SNP. **(b) SuSiE with CARMA vs. SLALOM.** Performance of credible sets ($\rho = 0.99$). Power: the proportion of simulated causal variants included in any credible set; Coverage: the proportion of credible sets containing a causal variant; Size: the number of variants included in a credible set; Purity: the mean squared correlation of variants in a credible set. The true number of causal variants is $|T| = 2$.

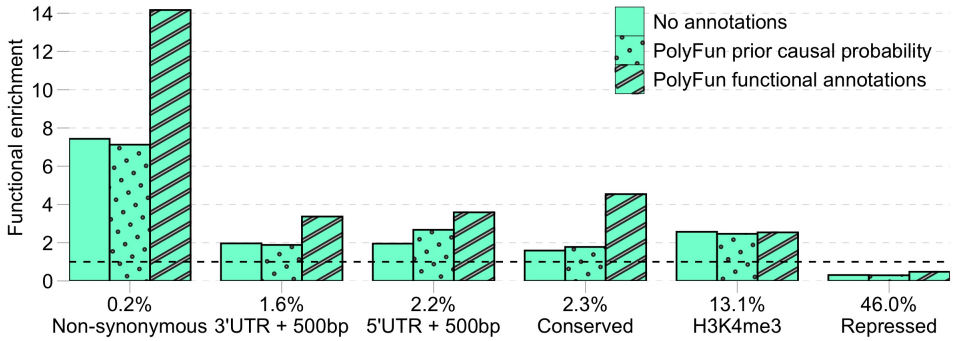

Figure 5: **Functional enrichments of prioritized SNPs at AD GWAS loci.** Functional enrichment of SNPs in credible models prioritized by CARMA, CARMA +PolyFun prior causal probability, and CARMA+Polyfun annotations in AD meta-analysis. The proportion of common SNPs lying in each functional category is reported above its name.

	AD loci											
	CR1	BIN1	CLNK	HS3ST1	CD2AP	ZCWPW1	MS4A6A	PICALM	ADAM10	SCIMP	ABI3	ABCA7
SuSiE on original AD datasets												
# PIPs > 0.99	3	8	5	7	5	7	5	5	5	5	4	9
# credible sets	4	9	8	7	6	10	5	5	6	7	8	10
SuSiE on filtered AD datasets based on CARMA outlier detection												
# PIPs > 0.99	0	1	0	0	0	0	0	0	0	0	0	1
# credible sets	1	1	1	1	1	1	1	1	1	0	0	2

Table 1: **SuSiE on the original AD data and filtered AD data based on CARMA outliers**. For each locus, the number of SNPs with PIP larger than 0.99 and the number of credible sets before and after CARMA outlier filtering are shown.

Locus	SNP	NoAnnot.	PolyFunProb.	PolyFunAnnot.	Z	N	EAF	CADD	EigenPC	Eigen	V2G	meQTL	eQTL
ALPK2	rs76726049	0.961	0.961	1.000	5.521	433227	0.011	3.554	2.938	4.090	*MALT1*		
CASS4	rs6014724	0.292	0.268	1.000	-6.208	433858	0.095	9.371	8.716	9.369	*CASS4*	✓	✓
SORL1	rs74685827	0.009	0.010	1.000	3.919	436498	0.018	6.325	17.254	4.850	*SORL1*	✓	
APH1B	rs117618017	0.485	0.442	1.000	5.518	444006	0.125	9.705	0.602	0.663	*APH1B*	✓	✓
BIN1	rs4663105	1.000	1.000	1.000	14.005	425134	0.411	0.348	0.904	0.026	*BIN1*	✓	✓
KAT8	rs11865499	0.099	0.113	1.000	-5.499	435397	0.299	4.182	4.315	5.328	*KAT8*	✓	✓
ABCA7	rs3752241	0.970	0.970	1.000	-6.279	428855	0.169	9.364	14.957	16.883	*ABCA7*	✓	✓
INPPD5	rs10933431	0.327	0.339	1.000	-6.153	436498	0.244	2.817	4.214	3.631	*INPP5D*	✓	
SLC24A4	rs11623883	0.260	0.233	0.998	-5.160	434845	0.466	0.68	0.287	0.369	*SLC24A4*	✓	✓
ABI3	rs616338	0.156	0.161	0.995	4.940	364859	0.006	15.9	0.424	0.240	*ABI3*		
ADAMTS4	rs4575098	0.716	0.716	0.994	6.370	434326	0.228	7.876	14.425	12.555	*NDUFS2*	✓	✓
TREM2	rs143332484	0.965	0.965	0.991	5.540	382336	0.007	11.42	0.744	0.630	*TREM2*		
MS4A6A	rs180896475	0.089	0.062	0.990	4.583	363900	0.003	9.641	4.134	4.850	*CBLIF*		
ADAM10	rs77145198	0.008	0.008	0.987	-5.059	443608	0.065	9.019	8.361	15.064	*SLTM*	✓	
BZRAP1-AS1	rs2526380	0.902	0.902	0.933	-5.568	71639	0.445	0.264	8.529	0.269	*MPO*	✓	
HESX1	rs184384746	0.826	0.826	0.905	5.695	364480	0.001	5.481	1.940	2.055	*IL17RD*		
EPHA1	rs3935067	0.421	0.051	0.903	6.562	426960	0.350	7.186	6.709	6.781	*EPHA1*	✓	
CLU-PTK2B	rs28834970	0.347	0.305	0.867	6.830	434770	0.366	5.444	6.171	4.850	*PTK2B*	✓	✓
PICALM	rs10792832	0.073	0.074	0.764	-8.666	436237	0.359	4.29	8.520	4.850	*PICALM*	✓	
SCIMP	rs75511804	0.097	0.101	0.697	6.049	435747	0.120	2.753	22.211	12.297	*RABEP1*	✓	✓
HS3ST1	rs13134197	0.241	0.274	0.632	5.439	71639	0.291	4.291	0.823	1.091	*HS3ST1*		
ECHDC3	rs11257238	0.000	0.000	0.534	5.724	430330	0.361	4.933	3.247	4.850	*USP6NL*		
CD33	rs3865444	0.430	0.432	0.498	-5.842	436498	0.299	0.345	7.971	1.550	*CD33*	✓	✓
ZCWPW1	rs2405442	0.282	0.279	0.472	-7.906	440597	0.323	8.402	16.605	14.470	*PILRB*	✓	✓
CD2AP	rs9381563	0.416	0.405	0.461	6.362	429204	0.356	0.27	12.254	2.614	*CD2AP*	✓	✓
CNTNAP2	rs116155464	0.217	0.220	0.310	5.954	364834	0.002	4.425	2.244	2.765	*CNTNAP2*		
CR1	rs6701713	0.173	0.157	0.298	8.837	443992	0.192	3.564	8.013	0.157	*CR1*		✓
CLNK	rs6448453	0.000	0.140	0.129	6.000	436498	0.262	3.22	3.222	5.329	*CLNK*	✓	

Table 2: **Top SNP (with largest PIP in CARMA with PolyFun Annotations) at each GWAS locus.** PIPs from CARMA without and with functional annotations (PolyFun prior probability or PolyFun annotations). Z: meta-analysis Z-score; N: number of samples; EAF: effect allele frequency; CADD/EigenPC/Eigen: PHRED-scaled scores for CADD/EigenPC/Eigen; V2G: top scoring gene from Open Targets; meQTL: indicates if the top gene from V2G is pointed to by meQTL data from ROSMAP Brain Dorsolateral Prefrontal Cortex; eQTL: indicates if the top gene from V2G is an eGene GTEx in a brain or blood tissue in GTEx.

Method	Prior on effect size	Prior on model space	Functional annot	Computation	Ref.
FINEMAP	Mixture of Normal -Gamma	discrete Uniform or discrete probability	Yes	Shotgun	[10]
CAVIARBF	Normal-Gamma	discrete probability	Yes	Exhaustive	[37]
JAM	g-prior	Beta-Binomial	No	Exhaustive and MCMC	[9]
fastPAINTOR	Normal-Gamma	discrete probability	Yes	Exhaustive and MCMC	[6]
SuSiE (+PolyFun)	Normal-Gamma	Multinomial (discrete probability)	Yes	Variational Bayes	[5, 12]
CARMA (proposed model)	Spike-and-slab/Cauchy	Poisson	Yes	Shotgun	

Table 3: **Summary of commonly used Bayesian fine-mapping methods**.