Bioinformatics, 2025, **41(1)**, btaf017
https://doi.org/10.1093/bioinformatics/btaf017
Advance Access Publication Date: 12 January 2025
Original Paper

OXFORD

Genetics and population analysis

Funmap: integrating high-dimensional functional annotations to improve fine-mapping

Yuekai Li[1], Jiashun Xiao[2], Jingsi Ming[3], Yicheng Zeng[2,*], Mingxuan Cai[1,*]

[1]Department of Biostatistics, City University of Hong Kong, Hong Kong, China
[2]Shenzhen International Center for Industrial and Applied Mathematics, Shenzhen Research Institute of Big Data, Shenzhen 518172, China
[3]Academy of Statistics and Interdisciplinary Sciences, KLATASDS-MOE, East China Normal University, Shanghai 200062, China

*Corresponding authors. Yicheng Zeng, Shenzhen International Center for Industrial and Applied Mathematics, Shenzhen Research Institute of Big Data, 2001 Longxiang Boulevard, Shenzhen 518172, China. E-mail: statzyc@gmail.com; Mingxuan Cai, Department of Biostatistics, City University of Hong Kong, 83 Tat Chee Avenue, Hong Kong, China. E-mail: mingxcai@cityu.edu.hk.

Associate Editor: Russell Schwartz

Abstract

Motivation: Fine-mapping aims to prioritize causal variants underlying complex traits by accounting for the linkage disequilibrium of genome-wide association study risk locus. The expanding resources of functional annotations serve as auxiliary evidence to improve the power of fine-mapping. However, existing fine-mapping methods tend to generate many false positive results when integrating a large number of annotations.

Results: In this study, we propose a unified method to integrate high-dimensional functional annotations with fine-mapping (Funmap). Funmap can effectively improve the power of fine-mapping by borrowing information from hundreds of functional annotations. Meanwhile, it relates the annotation to the causal probability with a random effects model that avoids the over-fitting issue, thereby producing a well-controlled false positive rate. Paired with a fast algorithm, Funmap enables scalable integration of a large number of annotations to facilitate prioritizing multiple causal single nucleotide polymorphisms. Our comprehensive simulations across a wide range of annotation relevance settings demonstrate that Funmap is the only method that produces well-calibrated false discovery rate under the setting of high-dimensional annotations while achieving better or comparable power gains as compared to existing methods. By integrating genome-wide association studies of 4 lipid traits with 187 functional annotations, Funmap consistently identified more variants that can be replicated in an independent cohort, achieving 15.5%–26.2% improvement over the runner-up in terms of replication rate.

Availability and implementation: The Funmap software and all analysis code are available at https://github.com/LeeHITsz/Funmap.

1 Introduction

Genome-wide association studies (GWASs) have successfully identified hundreds of thousands of single nucleotide polymorphisms (SNPs) that are statistically associated with complex human traits (Uffelmann *et al.* 2021). However, because the true causal variants may be correlated with many non-causal variants located in the proximal region (Slatkin 2008) due to linkage disequilibrium (LD), it remains very difficult to distinguish the causal SNPs from non-causal ones in GWAS discoveries. To address this issue, fine-mapping (Schaid *et al.* 2018) is employed to prioritize the set of variants that are most likely to be biologically associated with a target trait while accounting for the local LD pattern within each genetic locus identified by GWAS. By offering a set of putative causal SNPs that are potentially responsible for pathology of complex human diseases, fine-mapping outputs can serve as valuable resources for in-depth exploration to interpret disease etiology.

A number of fine-mapping methods have been developed to prioritize causal variants based on GWAS data. Early methods for fine-mapping (Hormozdiari *et al.* 2014, Kichaev *et al.* 2014) rely on an exhaustive search for combinations of causal SNPs, making them computationally inefficient in identifying multiple causal SNPs. Later, some methods employ approximate algorithms to reduce the computational cost of searching causal variants (Chen *et al.* 2015, Benner *et al.* 2016, Wen *et al.* 2016, Lee *et al.* 2018). More recently, SuSiE (Wang *et al.* 2020, Zou *et al.* 2022) proposes an efficient fine-mapping framework by decomposing causal signals into a sum of single causal effects. Despite the great advances in fine-mapping, it remains a major challenge to reliably prioritize causal SNPs when they are in strong LD with non-causal ones.

Fortunately, functional annotations can serve as auxiliary information to inform the prioritization of causal SNPs. This is granted by the fact that biologically important SNPs are more enriched within functionally important annotations across the genome. For example, a fine-mapping analysis across 14 traits reported that putative causal SNPs were significantly enriched in functional regions, including nonsynonymous regions, conserved regions, and multiple cell-type-specific regulatory regions (Weissbrod *et al.* 2020).

An expanding amount of annotation resources is becoming available to facilitate genetic studies. The Encyclopedia of DNA Elements (ENCODE) project (Consortium *et al.* 2012) has generated a comprehensive mapping between functional elements and variants covering 80% of the genome, including

Received: 23 June 2024; **Revised:** 20 December 2024; **Editorial Decision:** 7 January 2025; **Accepted:** 9 January 2025

open chromatin sites, histone mark enriched regions, and transcription factor binding regions, etc. The NIH Roadmap Epigenomics Mapping Consortium (Kundaje *et al.* 2015) has established high-quality, genome-wide maps of epigenomic regulatory elements across hundreds of human cell types and tissues, such as key histone modifications, chromatin accessibility, DNA methylation, and mRNA expression. Although the rich annotation resources hold promise to enhance fine-mapping, it requires handling high-dimensional annotation data, which hampers the effective integration of functional annotations. The difficulties are 2-fold. First, a large number of model parameters is required to characterize the relationship between annotations and the causal status of candidate SNPs. These parameters need to be estimated with GWAS data of only a few thousand correlated SNPs from a single genomic region, making the parameter estimation highly unreliable, thereby producing false positive results. Second, when the number of annotations grows large, it is computationally challenging to simultaneously search for multiple causal variants and estimate model parameters of high-dimensional annotations.

Much effort has been devoted to integrating functional annotations with GWAS data for fine-mapping. fastPAINTOR (Kichaev *et al.* 2017) can integrate multiple functional annotations with summary statistics by leveraging the approach of Markov chain Monte Carlo (MCMC). However, due to the lack of feature selection to filter out unrelated annotations, fastPAINTOR may incur a high false discovery rate (FDR) when dealing with high-dimensional functional annotations. To address this issue, PolyFun (Weissbrod *et al.* 2020) proposes to first estimate prior weights based on functional annotations and GWAS summary statistics, and then perform other fine-mapping methods with these estimated prior weights. Because the two-step design of PolyFun does not maximize the joint-likelihood function, it usually has suboptimal statistical power. CARMA (Yang *et al.* 2023) attempts to improve existing methods by incorporating high-dimensional functional annotations via a penalized logistic regression, allowing for more reliable integration of summary data and high-dimensional functional annotations. However, CARMA's algorithm alternates between a sampling procedure to explore the posterior distributions and fitting an elastic net to estimate annotation weights that involve a cross-validation step, making its time complexity relatively high as the number of candidate SNPs increases. Very recently, SparsePro (Zhang *et al.* 2023) extends SuSiE by allowing the prior causal probabilities to be linked to binary functional annotations. Nevertheless, similar to PolyFun, the process of computing prior weights and fitting the SuSiE model in SparsePro is separated, and the functional annotations used in SparsePro are restricted to binary types.

In this article, we propose a unified method to integrate high-dimensional *fun*ctional annotations with GWAS data to improve fine-*map*ping (Funmap). The success of Funmap relies on its three unique features. First, it effectively improves the power of fine-mapping by fully utilizing the information of high-dimensional functional annotations. Second, it relates the high-dimensional functional annotation to the prior causal probability with a random effects model that avoids the over-fitting issue, thereby producing a well-controlled FDR. Third, paired with a fast variational Bayes algorithm, Funmap enables scalable integration of a large number of annotations to facilitate prioritizing multiple causal SNPs. With comprehensive simulation studies, we show that Funmap is the only method that produces well-calibrated FDR under the setting of high-dimensional annotations while achieving better or comparable power gains as compared to existing fine-mapping methods, especially when only a subset of annotations is informative. We applied Funmap to prioritize causal SNPs of 4 lipid traits by integrating their GWAS data with 187 functional annotations. Our results suggest that Funmap not only boosts the statistical power by fully leveraging the auxiliary evidence of function annotations but also substantially improves the reproducibility of putative causal SNPs, indicating its effectiveness in reducing false positives. Furthermore, we demonstrate that Funmap's unique annotation importance measure provides valuable insights on how the functional annotations facilitate the mapping of causal variants, offering a deeper understanding of biological mechanisms underlying complex phenotypes.

2 Materials and methods

2.1 The Funmap model

We begin the formulation of Funmap with the individual level GWAS data. Consider the GWAS dataset $\{\mathbf{X}, \mathbf{y}\}$, where $\mathbf{X} \in \mathbb{R}^{n\times p}$ is the genotype matrix of n individuals and p SNPs of the target region, and $\mathbf{y} \in \mathbb{R}^n$ is the phenotype vector. Without loss of generality, we assume that each column of $\mathbf{X}$ and $\mathbf{y}$ has been standardized to have zero mean and unit variance. We also assume that the covariates, such as gender, age, and genotype principal components, have been properly adjusted following our previous works (Cai *et al.* 2021). We relate the phenotype $\mathbf{y}$ to genotypes $\mathbf{X}$ with the following linear model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}, \tag{1}$$

where $\mathbf{b} \in \mathbb{R}^p$ is the sparse vector of SNP effect sizes, the independent noise $\mathbf{e} \sim \mathcal{N}(0, \sigma^2\mathbf{I}_n)$, and $\mathbf{I}_n$ is the n by n identity matrix. To identify the non-zero entries of $\mathbf{b}$, we consider the following sum-of-single-effects (Wang *et al.* 2020) structure:

$$\mathbf{b} = \sum_{l=1}^{L} \boldsymbol{\gamma}_l b_l, \quad \boldsymbol{\gamma}_l \sim \text{Mult}(1, \boldsymbol{\pi}_l), \quad b_l \sim \mathcal{N}(0, \sigma_{bl}^2), \tag{2}$$

where $\boldsymbol{\gamma}_l = [\gamma_{l1}, \ldots, \gamma_{lp}]^T \in \{0,1\}^p$ is a binary vector with $\gamma_{lj} = 1$ indicating the lth causal signal is attributed to the jth SNP, b_l is the effect size of the lth causal signal, and $\boldsymbol{\pi}_l = [\pi_{l1}, \ldots, \pi_{lp}]^T$ is the vector of prior causal probabilities with $\sum_{j=1}^{p} \pi_{lj} = 1$.

Suppose that besides the GWAS data, we also collected the functional annotations of the target SNPs. Let $\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_p]^T \in \mathbb{R}^{p\times m}$ be the matrix collecting m annotations of the p SNPs. With the availability of rich functional annotations, m is becoming increasingly large (e.g. $m \geq 100$). To effectively incorporate the functional annotation as auxiliary evidence for causal SNP prioritization, we link the prior probability $\boldsymbol{\pi}_l$ to SNP annotations with the following softmax model with random effects:

$$\pi_{lj} = \frac{e^{\mathbf{A}_j^T \mathbf{w}_l}}{\sum_{j'=1}^{p} e^{\mathbf{A}_{j'}^T \mathbf{w}_l}}, \quad \mathbf{w}_l \sim \mathcal{N}(0, \sigma_{\mathbf{w}l}^2 \mathbf{I}_m), \tag{3}$$

where $\mathbf{w}_l \in \mathbb{R}^m$ represents the random effects vector of annotations on the causal probability of the lth single-effect component. Unlike previous methods (Kichaev *et al.* 2017, Zhang

et al. 2023) that assume a fixed annotation effect shared across single-effect components, we propose a component-specific random effects assumption on the annotation weights $\mathbf{w}_l$. This assumption has two salient properties. First, it allows the annotation weights to vary across causal signals, which better characterizes the real genetic architecture. Second, built upon a random effects model, when dealing with high-dimensional functional annotation $\mathbf{A}$, Funmap can adaptively estimate the annotation weights $\mathbf{w}_l$ from the data while shrinking those of redundant annotations to avoid over-fitting, thereby effectively reducing false positive results.

Under Equations (1–3), we denote the model parameters as $\boldsymbol{\theta} := \{\sigma^2, \sigma^2_{b1}, \ldots, \sigma^2_{bL}, \sigma^2_{w1}, \ldots, \sigma^2_{wL}\}$ and the collections of random variables $\tilde{\mathbf{b}} := \{b_1, \ldots, b_L\}$, $\tilde{\boldsymbol{\gamma}} := \{\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_L\}$, and $\tilde{\mathbf{w}} := \{\mathbf{w}_1, \ldots, \mathbf{w}_L\}$. The logarithm of the marginal likelihood is given as $\log \Pr(\mathbf{y}|\mathbf{X}, \mathbf{A}; \boldsymbol{\theta}) = \log \sum_{\tilde{\boldsymbol{\gamma}}} \int_{\tilde{\mathbf{b}}} \int_{\tilde{\mathbf{w}}} \Pr(\mathbf{y}, \tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}}|\mathbf{X}, \mathbf{A}; \boldsymbol{\theta}) d\tilde{\mathbf{w}} d\tilde{\mathbf{b}}$. By maximizing the log-likelihood, we aim to obtain parameter estimates $\widehat{\boldsymbol{\theta}}$ and prioritize the causal SNPs by using the posterior probability:

$$\Pr\left(\tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{b}}, \tilde{\mathbf{w}}|\mathbf{y}, \mathbf{X}, \mathbf{A}; \widehat{\boldsymbol{\theta}}\right) = \frac{\Pr\left(\tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{b}}, \tilde{\mathbf{w}}, \mathbf{y}|\mathbf{X}, \mathbf{A}; \widehat{\boldsymbol{\theta}}\right)}{\Pr\left(\mathbf{y}|\mathbf{X}, \mathbf{A}; \widehat{\boldsymbol{\theta}}\right)}. \quad (4)$$

2.2 Algorithm and parameter estimation

The major challenge to evaluating Equation (4) is the intractable marginal likelihood on the denominator. First, it is difficult to exactly evaluate the integration of $\Pr(\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}|\tilde{\mathbf{w}}, \mathbf{A}; \boldsymbol{\theta})$ due to the softmax function. Second, the sum-of-single-effects assumption makes it intractable to integrate over $\tilde{\boldsymbol{\gamma}}$. To address the above issues, we develop a variational inference algorithm that can efficiently estimate model parameters and approximate the posterior distribution.

We first deal with the intractability resulting from the softmax function. Here, we consider the following bound based on double majorization (Bouchard 2007):

$$\sum_{j'=1}^{p} e^{\mathbf{A}_{j'}^T \mathbf{w}_l} \leq \exp[\rho_l + \sum_{j'=1}^{p} \frac{\mathbf{A}_{j'}^T \mathbf{w}_l - \rho_l - \xi_{lj'}}{2} + \lambda(\xi_{lj'})((\mathbf{A}_{j'}^T \mathbf{w}_l - \rho_l)^2 - \xi_{lj'}^2) + \log(1 + e^{\xi_{lj'}})], \quad (5)$$

where $\lambda(\xi_{lj'}) = \frac{1}{2\xi_{lj'}}\left(\frac{1}{1+e^{-\xi_{lj'}}} - \frac{1}{2}\right)$, $\xi_{lj'} \in [0, \infty)$, and $\rho \in \mathbb{R}$. Clearly, the right-hand side of this inequality is in the exponentiated quadratic form. Therefore, applying this inequality to the denominator of the softmax function (3) leads to a tractable bound $g(\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}|\tilde{\mathbf{w}}, \mathbf{A}; \boldsymbol{\xi}, \boldsymbol{\rho}, \boldsymbol{\theta}) \leq \Pr(\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}|\tilde{\mathbf{w}}, \mathbf{A}; \boldsymbol{\theta})$, where $\boldsymbol{\xi} = [\xi]_{lj} \in [0, \infty)^{L \times p}$ and $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_L]^T \in \mathbb{R}^L$ are variational parameters that can be estimated from the data (see details in Supplementary Note S1). Let $\boldsymbol{\Theta} = \{\boldsymbol{\theta}, \boldsymbol{\xi}, \boldsymbol{\rho}\}$. We can obtain a lower bound of complete-data likelihood as:

$$\begin{aligned} \Pr(\mathbf{y}, \tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}}|\mathbf{X}, \mathbf{A}; \boldsymbol{\theta}) &\geq \Pr(\mathbf{y}|\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}}, \mathbf{X}; \boldsymbol{\theta}) g(\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}|\tilde{\mathbf{w}}, \mathbf{A}; \boldsymbol{\Theta}) \Pr(\tilde{\mathbf{w}}|\boldsymbol{\theta}) \\ &\equiv f(\mathbf{y}, \tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}}|\mathbf{X}, \mathbf{A}; \boldsymbol{\Theta}). \end{aligned} \quad (6)$$

Based on the above bound, we further derive a lower bound of the logarithm of marginal likelihood

$$\begin{aligned} &\log \Pr(\mathbf{y}|\mathbf{X}, \mathbf{A}; \boldsymbol{\theta}) = \log \textstyle\sum_{\tilde{\boldsymbol{\gamma}}} \int_{\tilde{\mathbf{b}}} \int_{\tilde{\mathbf{w}}} \Pr(\mathbf{y}, \tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}}|\mathbf{X}, \mathbf{A}; \boldsymbol{\theta}) d\tilde{\mathbf{w}}\, d\tilde{\mathbf{b}} \\ &\geq \log \textstyle\sum_{\tilde{\boldsymbol{\gamma}}} \int_{\tilde{\mathbf{b}}} \int_{\tilde{\mathbf{w}}} f(\mathbf{y}, \tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}}|\mathbf{X}, \mathbf{A}; \boldsymbol{\Theta}) d\tilde{\mathbf{w}}\, d\tilde{\mathbf{b}} \\ &\geq \textstyle\sum_{\tilde{\boldsymbol{\gamma}}} \int_{\tilde{\mathbf{b}}} \int_{\tilde{\mathbf{w}}} q(\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}}) \log \frac{f(\mathbf{y}, \tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}}|\mathbf{X}, \mathbf{A}; \boldsymbol{\Theta})}{q(\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}})} d\tilde{\mathbf{w}}\, d\tilde{\mathbf{b}} \\ &= \mathrm{E}_q[\log f(\mathbf{y}, \tilde{b}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}}|\mathbf{X}, \mathbf{A}; \boldsymbol{\Theta}) - \log q(\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}})] \equiv F(q, \boldsymbol{\Theta}), \end{aligned} \quad (7)$$

where the first inequality follows the double majorization bound, the second inequality is granted by Jensen's inequality, and $q(\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}})$ is an approximation of the exact posterior (4). To analytically evaluate the lower bound, we introduce the following factorizable mean-field formulation assumption:

$$q(\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}}) = \prod_{l=1}^{L} q_l(b_l, \boldsymbol{\gamma}_l, \mathbf{w}_l) = \prod_{l=1}^{L} q_l(b_l, \boldsymbol{\gamma}_l) q_l(\mathbf{w}_l), \quad (8)$$

where $q_l(b_l, \boldsymbol{\gamma}_l) = q(b_l|\boldsymbol{\gamma}_l) q(\boldsymbol{\gamma}_l)$. The above approximation can be obtained in closed form, which allows the lower bound to be analytically evaluated. Besides, by taking advantage of the sum-of-single-effects decomposition, this mean-field assumption inherits the property of SuSiE that only requires the L causal signals to be independent, relaxing the assumptions in traditional variational approximations (Guan and Stephens 2011, Carbonetto and Stephens 2012). We develop a variational inference algorithm that can efficiently estimate $\boldsymbol{\theta}$ and evaluate $q(\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}}, \tilde{\mathbf{w}})$ by iteratively maximizing the lower bound (7) under assumption (8).

To improve the convergence of the variational algorithm, we design a three-stage model fitting process with warm-starts. This is built upon the knowledge that Funmap covers the SuSiE model as a special case when all the annotation effects $\tilde{\mathbf{w}} = 0$. In the first stage, we fit the SuSiE model using our Funmap implementations without incorporating annotations, obtaining an initial estimate of parameters $\sigma^2, \sigma^2_{b1}, \ldots, \sigma^2_{bL}$ and the posterior distribution $q(\tilde{\mathbf{b}}, \tilde{\boldsymbol{\gamma}})$. Then, we use these estimates to initialize the second stage where the variational algorithm is executed to evaluate the posterior distributions of $\tilde{\mathbf{w}}$. In this stage, we fix $\tilde{\boldsymbol{\gamma}}$ at its posterior mean obtained in the first stage to produce rough estimates of annotation-related parameters σ^2_{wl} and $q(\tilde{\mathbf{w}})$. Finally, we use the estimated parameters from the second stage as the initial values to run the variational algorithm with the full Funmap model.

The Funmap fitting procedure possesses several appealing computational and practical convenience. Computationally, since the three fitting stages are constructed under a unified variational framework, the lower bound is guaranteed to increase at each stage, enabling a more stable convergence of the Funmap algorithm. Practically, it avoids the need to specify the number of causal variants. Specifically, when the number of causal SNPs is unknown, we can simply set L to a reasonably large number. The excessive components of Funmap will be assigned small posterior probabilities. This property can be attributed to two facts. Firstly, at the first stage, excessive components will be broadly assigned to all SNPs across the locus due to the high uncertainty in allocating their causal effects (Wang *et al.* 2020). Secondly and importantly, the warm-start procedure can prevent the excessive components from being enlarged by over-fitting the

high-dimensional annotations. We provide the details of our algorithm in the Supplementary Note S1.

2.3 Funmap with GWAS summary data

While the variational inference algorithm proposed above is for individual-level genotype $\mathbf{X}$ and phenotype $\mathbf{y}$, it can be easily extended to using only GWAS summary data as input. Let us consider the z-scores obtained from marginal regressions:

$$z_j = \frac{\widehat{\beta}_j}{\widehat{s}_j}, \text{ where } \widehat{\beta}_j = \left(\mathbf{x}_j^T\mathbf{x}_j\right)^{-1}\mathbf{x}_j^T\mathbf{y}, \widehat{s}_j = \sqrt{\frac{||\mathbf{y}-\mathbf{x}_j\widehat{\beta}_j||_2^2}{n\mathbf{x}_j^T\mathbf{x}_j}}, \quad (9)$$

and $\mathbf{x}_j \in \mathbb{R}^p$ is the jth column of $\mathbf{X}$. We first note that the likelihood and its lower bound (7) depend on the GWAS data $\{\mathbf{X}, \mathbf{y}\}$ only through the sufficient statistics $\mathbf{X}^T\mathbf{X}$, $\mathbf{X}^T\mathbf{y}$, and $\mathbf{y}^T\mathbf{y}$. Since both $\mathbf{X}$ and $\mathbf{y}$ have been standardized, we can replace the sufficient statistics with z-scores and LD matrix $\mathbf{R}$ (Wang *et al.* 2020) with the following relationships:

$$\mathbf{X}^T\mathbf{X} = n\mathbf{R}, \quad \mathbf{x}_j^T\mathbf{y} = \frac{n}{\sqrt{n+z_j^2}}z_j, \quad \mathbf{y}^T\mathbf{y} = n, \quad (10)$$

where $\mathbf{R}$ can be computed with genotypes from a subset of GWAS samples or from a reference panel of similar ancestry background. In our software implementation, all the operations are based on sufficient statistics, allowing us to accommodate both individual-level and summary-level GWAS data.

2.4 Identification of causal SNPs

After the convergence of Funmap algorithm, we can obtain the approximate component-specific posterior probabilities $q(\boldsymbol{\gamma}_l)$, where $q(\boldsymbol{\gamma}_{lj} = 1)$ represents the probability that jth SNP is responsible for the lth causal signal. With this information, we can compute the posterior inclusion probability of the jth SNP as $\text{PIP}_{(j)} = \Pr(\gamma_{lj} \neq 0 \text{ for some } l|\mathbf{X}, \mathbf{y}; \widehat{\boldsymbol{\theta}}) \approx 1 - \prod_{l=1}^{L}(1 - q(\boldsymbol{\gamma}_{lj} = 1))$. Then, the local FDR for each SNP can be calculated as $\text{fdr}_{(j)} = 1 - \text{PIP}_{(j)}, j = 1, \ldots, p$. To control the global, we sort the local FDRs in ascending order. We denote the ith sorted local FDR as $\text{fdr}_{(i)}$ and compute the global FDR as $\text{FDR}_{(j)} = \frac{\sum_{i=1}^{j}\text{fdr}_{(i)}}{j}$. If $\text{FDR}_{(j)}$ is lower than a given FDR threshold η, we consider the SNP as a putative causal SNP.

2.5 Importance score of functional annotations

To evaluate the relative importance of functional annotations in facilitating fine-mapping, we define a feature importance score based on the posterior distribution of the annotation weights $\mathbf{w}_l$. For the jth annotation, its feature importance FIj is defined as:

$$\text{FI}_j = \max_l \mu_{wlj}^2, \quad (11)$$

where μ_{wlj} is the posterior mean of the jth annotation weight in the lth component (see details in Equation (26) of the Supplementary Note S1). This definition captures the strongest evidence of annotation relevance across all components, as each component may have different relevant functional annotations under Funmap's component-specific random effects design. Based on these importance scores, annotations can be ranked and prioritized, with top-ranked ones (e.g. top 5% or 10%) considered potentially relevant.

3 Results

3.1 Simulation

We conducted comprehensive simulations to evaluate the performance of Funmap using real genotypes obtained from the UK Biobank (UKBB) project (Bycroft *et al.* 2018). First, we obtained genotypes of UKBB White British individuals with sample sizes $n = 20\,000$ and $50\,000$. To mimic the realistic genetic architectures of disease-related loci, we considered 10 risk regions identified in GWAS as associated with breast cancer (Fachal *et al.* 2020), where each region comprises approximately (Fachal *et al.* 2020), where each region comprises approximately $p = 700 \sim 4200$ variants (Supplementary Table S1). The functional annotations A_{jk} of these SNPs were sampled from a standard normal distribution for $j = 1, \ldots, p$ and $k = 1, \ldots, m$. We varied the number of annotations m among $\{20, 50, 100\}$ to cover the cases from low to high dimensional settings. Then, we constructed the effects of functional annotations $\mathbf{w} \in \mathbb{R}^m$ with $w_k \sim \mathcal{N}(0, 0.01)$ for $k = 1, \ldots, m$, where each annotation only has a small impact on the causal probability. Next, we simulated the causal probabilities with the softmax function $\pi_j = \frac{e^{\mathbf{A}_j^T\mathbf{w}}}{\sum_{j'=1}^{p} e^{\mathbf{A}_{j'}^T\mathbf{w}}}$. We selected L_0 SNPs as the true causal SNP which has the highest values of π_j and pairwise correlation less than 0.1 following previous studies (Yang *et al.* 2023, Zhang *et al.* 2023). We considered $L_0 = 2$ and $L_0 = 3$ to focus on scenarios where a locus harbours multiple causal variants. Given the causal status, we followed previous studies (Wang *et al.* 2020) to sample the effect sizes of causal SNPs $\mathbf{b}_0 \in \mathbb{R}^{L_0}$ from $\mathcal{N}(\mathbf{0}, 0.0075/L_0 \cdot \mathbf{I}_{L_0})$, where 0.0075 is the heritability contributed by the target region. With the standardized genotype matrix of causal SNPs $\mathbf{X}_0 \in \mathbb{R}^{n \times L_0}$, we generated the phenotype values with $\mathbf{y} = \mathbf{X}_0\mathbf{b}_0 + \mathbf{e}$, where $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \frac{1-0.0075}{0.0075}\text{Var}(\mathbf{X}\mathbf{b}_0) \cdot \mathbf{I}_n)$. Finally, we performed univariate linear regressions on each SNP to obtain their z-scores. For each setting, we repeated the experiment for 50 times across the 10 regions, yielding 500 replicates. All the setting parameters and results are in Supplementary Figs S1–S30.

In our simulation analysis, we benchmarked the performance of Funmap by comparing it with five representative fine-mapping methods. For methods that do not use functional information, we considered SuSiE (Wang *et al.* 2020), CARMA (Yang *et al.* 2023), and PAINTOR (Kichaev *et al.* 2017). For methods that use functional information, we considered the extensions of CARMA and PAINTOR for integrating functional annotations, denoted as "CARMA+anno" (Yang *et al.* 2023) and "PAINTOR+anno" (Kichaev *et al.* 2017), respectively. The number of components L for SuSiE and Funmap was set to 10, with no more than 100 iterations per stage. The expected number of causal SNPs (η) for CARMA and CARMA+anno was also set to 10, and the maximum number of iterations was set to 10 for both outer and inner loops. For PAINTOR and PAINTOR+anno, we used the fastPAINTOR implemented in the PAINTOR software v3.0 by adding the "-mcmc" flag.

We first assessed the FDR calibration of the compared methods. Specifically, we calculated the global FDR defined

and identified a set of putative causal SNPs with an expected FDR threshold. Given the putative causal SNPs, the empirical FDR can be computed as $\text{eFDR} = 1 - \frac{\text{NO. of true causal SNPs among putative causal SNPs}}{\text{NO. of putative causal SNPs}}$. The eFDR of a well-calibrated fine-mapping method should be close to the expected FDR level. Figure 1a and b show the eFDR against the expected FDR when $n = 50\,000$ and $L_0 = 2$. Clearly, all methods without annotation had well-controlled FDR across different thresholds. Funmap consistently achieved satisfactory performance across different numbers of annotations, suggesting its ability to handle high-dimensional annotations through the random effects model. By contrast, due to the fixed effect assumption on annotation weights, PAINTOR+anno exhibited severely inflated FDR, which was exacerbated as the number of annotations increased. Although CARMA+anno could also regularize the annotation weights by fitting an elastic net, its FDR was still inflated when there was a large number of annotations ($m \geq 50$). This may be attributed to the lack of a warm-start procedure in CARMA's algorithm, which leads to local optimal solutions. To illustrate the differences between the performance of compared methods more intuitively, we visualized the results of one dataset from our simulation experiments with Manhattan plots (Fig. 1c). We can see that in this example, the two causal SNPs are much more significant than those of the surrounding SNPs due to the weak LD. Given this evidence, it is relatively easy to prioritize the causal SNPs in this dataset. Indeed, the causal SNPs could be reliably identified

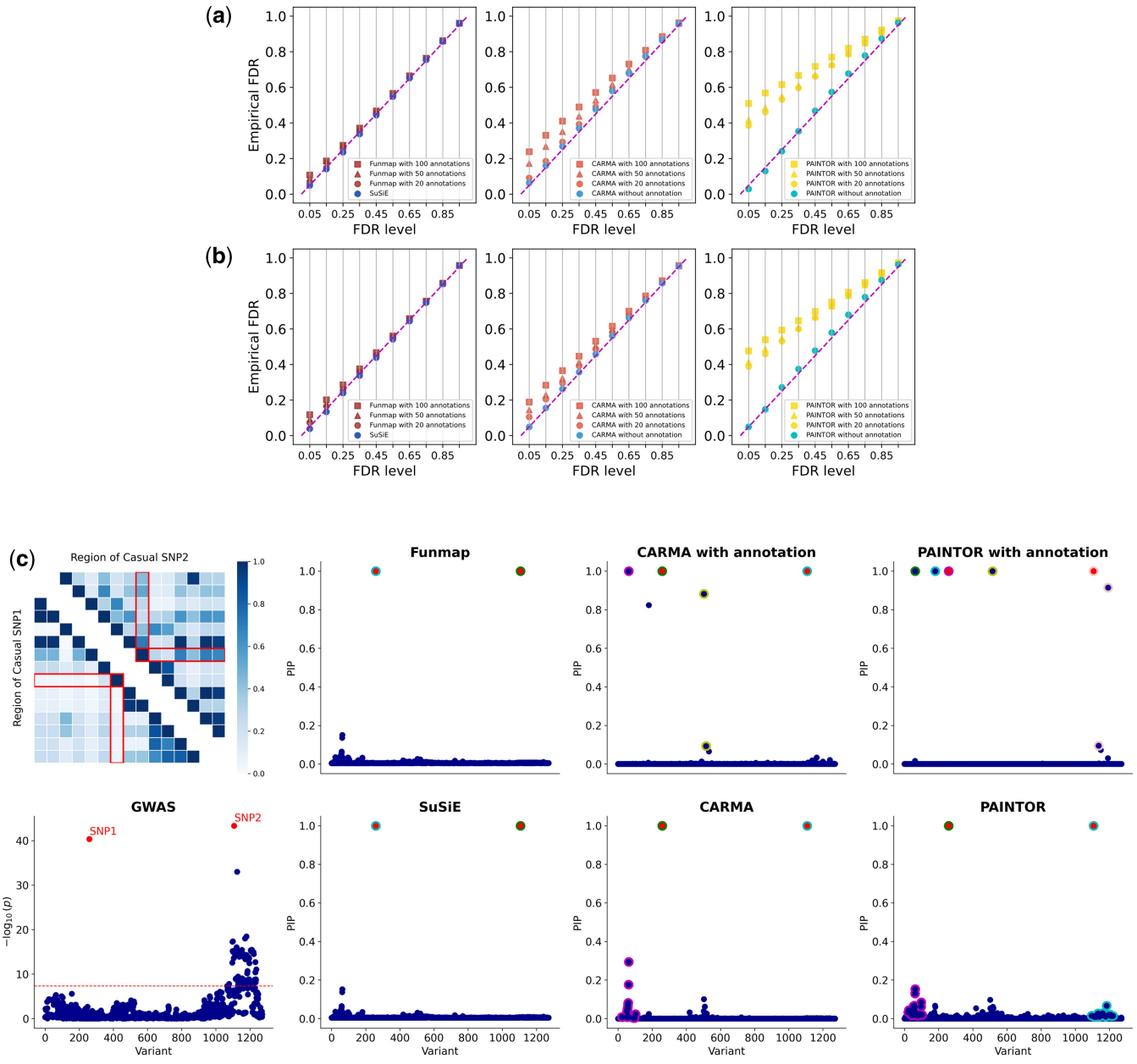

Figure 1. Comparison of FDR control in simulation studies. (a, b) Calibration of FDR with $n = 50\,000$, $m = 100$, while the number of causal SNPs is set to $L_0 = 2$ (a) and $L_0 = 3$ (b). Results are summarized from 500 replications across 10 regions. (c). An illustrative example generated by simulation. The first column shows the absolute correlation among the two candidate causal SNPs and their neighboring SNPs and the Manhattan plot. The second to fourth columns show the PIP obtained by with compared methods. Red dots represent causal SNPs. Dots with the same color of outline represent SNPs in the level-95% credible sets of a causal signal.

without the functional information. However, after incorporating functional annotations, CARMA+anno and PAINTOR+anno produced a number of false positive results. PINTOR+anno created four false positive signals with PIP >0.9. CARMA+anno also incorrectly assigns a high PIP (>0.8) to three non-causal SNPs. By contrast, Funmap effectively avoided false positive findings and yielded results similar to SuSiE. This observation implies the effectiveness of using Funmap's random effects model to adaptively incorporate functional information in high-dimensional settings.

Next, we considered a set of common PIP thresholds $\{0.90, 0.95, 0.99\}$ to evaluate the statistical power of compared methods. As shown in Fig. 2a and b, by incorporating functional annotation information, the statistical power of Funmap, CARMA+anno, and PAINTOR+anno outperformed their counterparts without annotations. The statistical power of Funmap was higher than CARMA+anno when stringent PIP thresholds (PIP > 0.99 and PIP > 0.95) were applied and comparable to CARMA+anno under a less stringent PIP threshold (PIP4 > 0.9). Although PAINTOR+anno produced the highest power in most settings, it is worth noting that its FDR can be strongly inflated when the number of annotations is large, making it difficult to replicate and interpret the discoveries. A concrete example with $m = 100$ and $L_0 = 2$ is given in Fig. 2c, where the two causal SNPs are in strong LD with neighboring SNPs. SuSiE and CARMA produced a PIP of around 0.8 for the causal SNPs, while PAINTOR's PIP failed to exceed 0.4. Despite the strong LD,

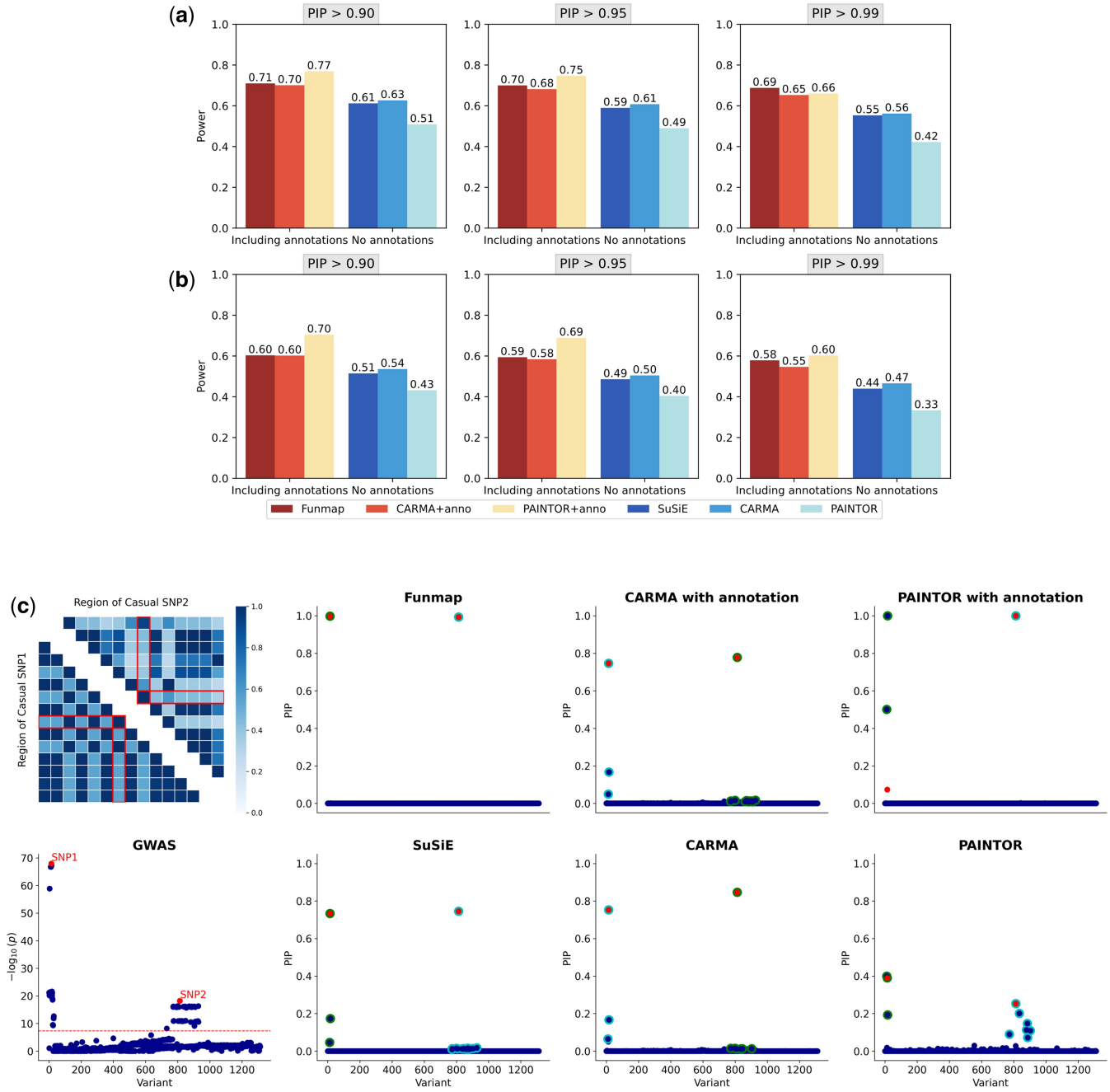

Figure 2. Comparison of statistical power in simulation studies. (a, b) Statistical power of compared methods with $n = 50\,000$, $m = 100$ while the number of causal SNPs is set to $L_0 = 2$ (a) and $L_0 = 3$ (b). (c) An illustrative example generated by simulation.

Funmap successfully elevated the PIP of causal SNPs to 1.0 by utilizing the functional information. In contrast, CARMA+anno remained largely unchanged and PAINTOR+anno incorrectly assigned high PIP to a non-causal SNP, yielding a false positive result. This example indicates that even in the presence of strong LD, traditional methods can have limited power, whereas Funmap can still effectively improve statistical power by leveraging the functional information while controlling FDR. To gain further insight into the difference in fine-mapping performance, we contrasted the PIP obtained by Funmap against those obtained by other methods (Fig. 3a and b). It is clear that the PIP of causal variants produced by Funmap was substantially larger than those produced by SuSiE and CARMA+anno, indicating Funmap's better ability to pinpoint the causal SNPs. As a comparison, PAINTOR+anno identified more causal SNPs with a cost of producing many false positives. This is in line with our observation of inflated FDR yielded by PAINTOR+anno. Besides the PIP for SNP-level inference, Funmap also provides a set of candidate SNPs for each causal signal, denoted as "credible set." A level-δ credible set includes a subset of SNPs that are responsible for a single-effect component with a joint probability of δ. Given a coverage level δ, a smaller credible set indicates a lower uncertainty to assign the causal signal and hence higher resolution of fine-mapping. We summarized the size of level-95% credible sets in Fig. 3d. As expected, Funmap yielded the smallest credible sets, suggesting its higher resolution.

To evaluate the computational efficiency of our model, we benchmarked the CPU time of compared methods by varying the number of variables and the number of functional annotations. We first focused on loci with the number of SNPs

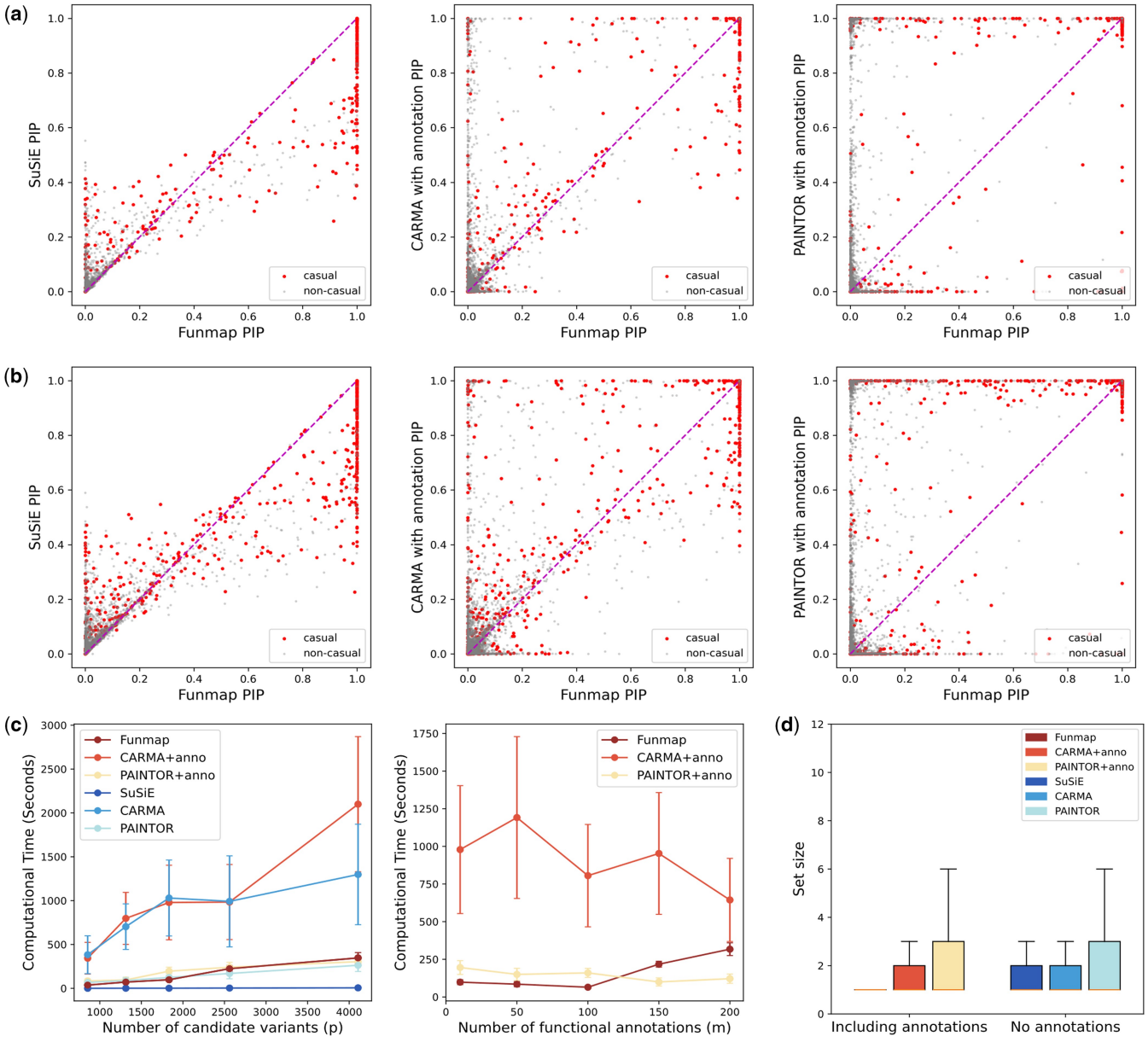

Figure 3. Comparison of PIP and CPU timings. (a, b). Comparison of PIP between Funmap and SuSiE (left panel), CARMA+anno (middle panel), and PAINTOR+anno (right panel) with $n = 50\,000$, $m = 100$, while L_0 is varied at 2 (a) and 3 (b). (c) CPU timings are shown for increasing p with $m = 100$ (left panel) and increasing m with $p = 1833$. (d) Boxplot displays the size of the 95% credible sets from the simulation results with $n = 50\,000, m = 100, L_0 \in \{2, 3\}$.

$p \in \{854, 1313, 1833, 2561, 4107\}$, while fixing the number of functional annotations at $m = 100$. Figure 3c shows the CPU time of all compared methods. As we can observe, Funmap achieved a great computational efficiency among methods capable of integrating functional annotation, with a CPU time substantially faster than CARMA+anno. While SuSiE was the fastest method, it could not integrate high-dimensional annotations to improve fine-mapping. The right panel of Fig. 3c compares the running times under different numbers of annotations $m \in \{10, 50, 100, 150, 200\}$ with the number of SNPs fixed at $p = 1833$. Overall, CARMA+anno had the largest computational overhead, with an average CPU time ranging between 10 and 16 minutes. Funmap achieved the least computing time when $m \leq 100$ and only had a marginal increase when m exceeds 150. Although PAINTOR+anno had CPU times comparable Funmap, it is not suitable to integrate high-dimensional annotations due to inflated FDR.

To assess the robustness of Funmap when our assumption of annotation effects is violated, we conducted a series of additional simulations that consider the sparse annotation weights. We varied the number of relevant annotations at 50, 10, 5, and 1 out of 100 annotations while keeping others irrelevant, corresponding to sparsity levels of 50%, 10%, 5%, and 1%, respectively. As shown in Supplementary Figs S31–S66, Funmap consistently produced better controlled FDR across all settings of sparsity levels, highlighting its robustness to the model mis-specification. In contrast, CARMA and PAINTOR had inflated the FDR regardless of the proportion of relevant annotations. Across all settings of annotation relevance, Funmap achieved better or comparable statistical power as compared to PAINTOR and CARMA. This pattern indicates that Funmap can adaptively integrate useful functional information to enhance statistical power while maintaining FDR control across a wide range of annotation relevance. In addition, to investigate whether Funmap correctly integrated the relevant annotations in the sparse setting, we compared the annotation selection performance Funmap with PAINTOR using their feature importance scores. CARMA was excluded from this analysis because its software does not provide annotation weight estimates or feature importance measures. We ranked annotations by importance scores and treated the top-ranked annotations as relevant with percentage thresholds of 5%, 10%, 20%, and 50%. The comparison was based on two metrics: statistical power, defined as the ratio of the number of relevant annotations included in the top-ranked group to the total number of relevant annotations, and the false positive rate (FPR), computed as the ratio of the number of relevant annotations included in the top-ranked group to the total number of irrelevant annotations. As shown in Supplementary Figs S67–S82, Funmap demonstrated superior performance, particularly in sparse settings where relevant annotations were rare (1% of total annotations). While both methods showed decreased power as the proportion of relevant annotations increased, Funmap consistently outperformed PAINTOR in terms of statistical power across all settings while maintaining comparable FPR, suggesting that Funmap can better distinguish relevant annotations from irrelevant ones to facilitate the identification of causal variants.

3.2 Real data analysis

We applied Funmap to identify causal SNPs for four lipid-related traits including high-density lipoprotein (HDL), low-density lipoprotein (LDL), triglycerides (TG), and total cholesterol (TC). The GWAS summary data were collected from 315 133 UKBB individuals of European ancestry. To construct candidate genomic regions for fine-mapping, we first identified genome-wide significant associations (P–values $< 5 \times 10^{-8}$) for each trait. Then, we extracted a 1-Mbp window centering at each genome-wide significant SNP. In total, we obtained 864 genomic regions (190–347 per trait) with 434–8646 SNPs per region across the four traits. To avoid spurious results as reported by previous studies (Weissbrod *et al.* 2020), we excluded the region of major histocompatibility complex (25.5 Mbp–33.5 Mbp in chromosome 6) and two regions with long-range LD (8 Mbp–12 Mbp in chromosome 8 and 46 Mbp–57 Mbp in chromosome 11). We followed previous studies (Weissbrod *et al.* 2020) to use 187 functional annotations from the Baseline-LF v2.2. UKB annotations (Gusev *et al.* 2014). Rather than being specifically focused on lipid-related functions, these annotations represent a systematic collection of genomic and functional features, encompassing a broad spectrum of functional elements, including coding regions, conserved elements, regulatory features, minor allele frequency (MAF) distributions, and LD-related characteristics, which have been demonstrated to be informative for various complex traits (Weissbrod *et al.* 2020). During the fine-mapping process, we considered one genomic region at a time, using the GWAS marginal z-scores, in-sample LD correlation matrices, and the functional annotations of local SNPs as inputs for all methods. We recorded the PIPs and credible sets for each genomic region. We used the same parameter settings as those used in simulation studies for Funmap, SuSiE, and PAINTOR. For CARMA, we set the maximum number of iterations to four for both inner and outer loops.

We summarized the putative causal SNPs in Supplementary Table S2. By integrating functional annotations, Funmap successfully identified more SNPs than SuSiE, PAINTOR, and CARMA across different PIP thresholds. Although PAINTOR+anno and CARMA+anno reported more causal SNPs than Funmap, our simulation analyses suggest that their discoveries may be unreliable due to the large number of annotations. To assess the credibility of our fine-mapping results, we extracted the set of SNPs that were reported as causal by each functionally informed method but not by SuSiE with a PIP threshold of 0.95. Then, we evaluated the replication rates of these newly identified SNPs using independent multi-ancestry GWASs from the Global Lipids Genetics Consortium (GLGC) (Graham *et al.* 2021), which comprised up to 900K samples across the four traits. The credible sets in this replication cohort were computed by SuSiE using meta-analyzed multi-ancestry data, yielding a highly reliable set of putative causal SNPs. Therefore, we can use the GLGC data as a high-quality resource to validate the fine-mapping results in our UKBB analysis. We evaluated the replication rates with two quantities: the proportion of new discoveries that were genome-wide significant (P-value $< 5 \times 10^{-8}$) and the proportion of new discoveries included in the credible sets in GLGC cohort. As summarized in Fig. 4, Funmap consistently yielded the highest replication rates across the four traits, with >50% of new discoveries successfully replicated in HDL, LDL, and TC in terms of genome-wide significance. Among the three methods that integrate functional annotations, PAINTOR+anno detected 1110 new associations with HDL that were not reported by SuSiE.

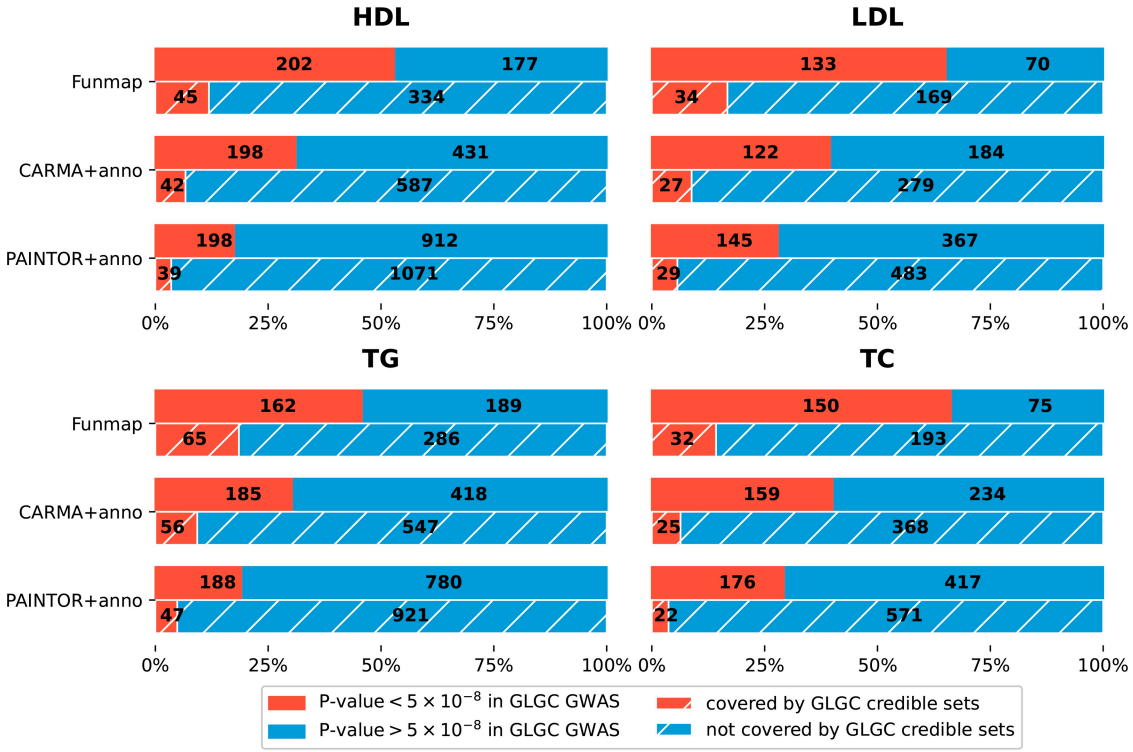

Figure 4. Replication analysis of Funmap, CARMA+anno, and PAINTOR+anno. Bar charts on the top shows the fraction and number of newly identified SNPs with *P*-value $<5\times10^{-8}$ in the replication cohorts of GLGC GWAS. Bar charts on the bottom shows the fraction and number of newly identified SNPs that are included in the 95%-level credible sets generated from GLGC GWAS with SuSiE.

However, only 17.8% (198/1110) were genome-wide significant, and 3.5% (39/1110) were included in the credible sets of GLGC cohort. Although CARMA+anno had better replication performance than PAINTOR+anno, only 31.5% (198/629) new discoveries were genome-wide significant in GLGC, meaning that more than half of its new discoveries could not be replicated. The low replication rates suggest that many new discoveries could be false positives due to the integration of high-dimensional annotations. By introducing the random effects model, Funmap substantially improved the replication rates. For example, in terms of genome-wide significance, 53.3% (202/379) new SNPs reported by Funmap were successfully replicated. Besides, 16.7% (34/203) of Funmap's new discoveries were included in the GLGC credible sets, which was 2.5 times higher than CARMA+anno and 4.8 times higher than PAINTOR+anno. It is worth noting that the number of Funmap's new discoveries replicated by GLGC credible sets was consistently higher, implying a greater statistical power. Meanwhile, the number of non-replicated SNPs was substantially reduced, indicating that many false positives were excluded. This pattern was consistent across the four traits. Therefore, the higher replication rates of Funmap can be attributed to both improved power and reduced false positives by properly integrating the high-dimensional annotations.

Funmap not only improved the replication rates but also enhanced the fine-mapping resolution in the lipid traits analysis. As observed in Fig. 5a, Funmap created the smallest credible sets among all compared methods, with a median size of one across the four traits. This observation was consistent with our simulation studies. As an example, we focus on the locus 6-7Mb in Chromosome 8, which harbours the SNP rs2928176 fine-mapped for TC only by Funmap (PIP = 1.0). This SNP locates on the 2Kb upstream of *AGPAT5*, encoding an integral membrane protein of the 1-acylglycerol-3-phosphate O-acyltransferase family (Strembitska *et al.* 2022). This enzyme is responsible for converting lysophosphatidic acid to phosphatidic acid, serving as a critical step of *de novo* phospholipid biosynthesis. The SNP correlation heatmap (top left panel in Fig. 5b) in this region shows that multiple SNPs around rs2928176 are in strong LD, making them highly significant with very similar *P*-values (bottom left panel in Fig. 5b). Therefore, traditional fine-mapping methods lacking the ability to integrate functional information could not reliably prioritize the causal SNP in this region. Indeed, SuSiE produced a very large 95% credible set comprising 41 SNPs, with the largest PIP only attaining 0.217. CARMA and PAINTOR generated credible sets of similar sizes. By incorporating the functional information, Funmap uniquely identified rs2928617 as the causal SNP with PIP = 1, producing a high-resolution credible set that only includes rs2928617. By contrast, CARMA+anno yielded three credible sets with sizes of one and two, which was suspicious as there is no strong evidence of three independent causal signals in this region. While PAINTOR+anno elevated the highest PIP to 0.8 and slightly reduced the size of the credible set, it failed to include rs2928617. This example consolidates our conclusion of Funmap's ability to improve power and resolution by incorporating high-dimensional functional annotations.

To gain insight into how and which annotations are contributing to the improved performance of fine-mapping, we conducted a comprehensive assessment of feature importance across all four lipid traits. We visualized the distribution of

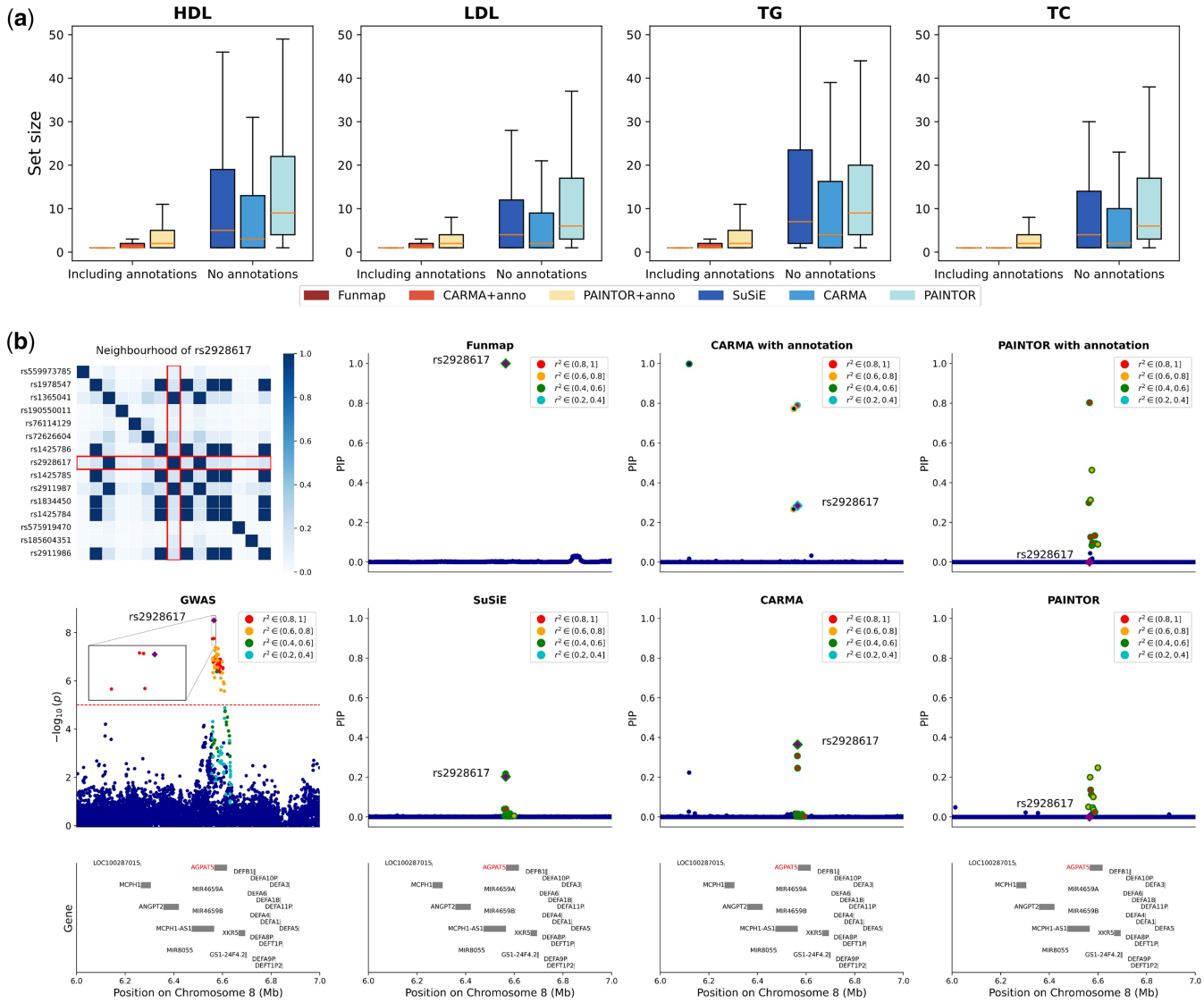

Figure 5. Comparison of credible set size and fine-mapping results from a region of TC GWAS. (a) Box plots of credible set size across four lipid traits. (b) Fine-mapping results of TC from locus 6 Mb–7 Mb in chromosome 8. The first column shows the heatmap of absolute correlation between rs2928617 and its neighboring SNPs and the Manhattan plot. The red dashed line represents 5×10^{-8}. The second to fourth column show the PIP obtained by with compared methods. The purple square represents SNP rs2928617 and the color of the points represents the correlation between neighboring SNPs and rs2928617. Dots with the same color of outline represent SNPs in the level-95% credible sets of a causal signal.

feature importance scores for each annotation across all causal signals through box plots. In Supplementary Figs S83–S86, we showed the 10 annotations with the highest median feature importance scores across all loci for HDL, LDL, TG, and TC, respectively. Our analysis demonstrated consistent patterns in the top-ranking functional annotations across all four lipid traits. Therefore, we summarized the distribution of feature importance scores across all four lipid traits in Fig. 6. Three categories of annotations emerged as particularly influential: nucleotide diversity, Genomic Evolutionary Rate Profiling number of substitution score (GERP-NS), and recombination rate. According to previous study, these annotations are fundamental in understanding the genetic architecture of complex human traits under the influence of negative selection (Gazal *et al.* 2017). Specifically, higher scores in nucleotide diversity, GERP-NS, and recombination rate were associated with lower probabilities of SNPs being causal, reflecting the impact of purifying selection on potentially deleterious variants. This pattern is particularly evident in regions with high recombination rates, where the Hill-Robertson effect leads to reduced heritability through more effective negative selection. Notably, all these annotations have a negative median effect on the causal probability, suggesting a higher annotation score indicates a lower probability of the SNPs to be causal. Indeed, the original study of the three annotations (Gazal *et al.* 2017) also reported the greater annotation score was associated with a lower heritability, suggesting the action of negative selection. This consistency indicates that Funmap successfully captures the impact of negative selection acting on causal signals and utilizes this information to facilitate fine-mapping of causal variants. Besides, consistent with previous enrichment analysis of causal variants in 16 UKBB traits that include lipid traits (Weissbrod *et al.* 2020), Funmap feature importance showed strong effects of super enhancer, transcribed regions, and histone marker acetylation of histone H3 at lysine 27

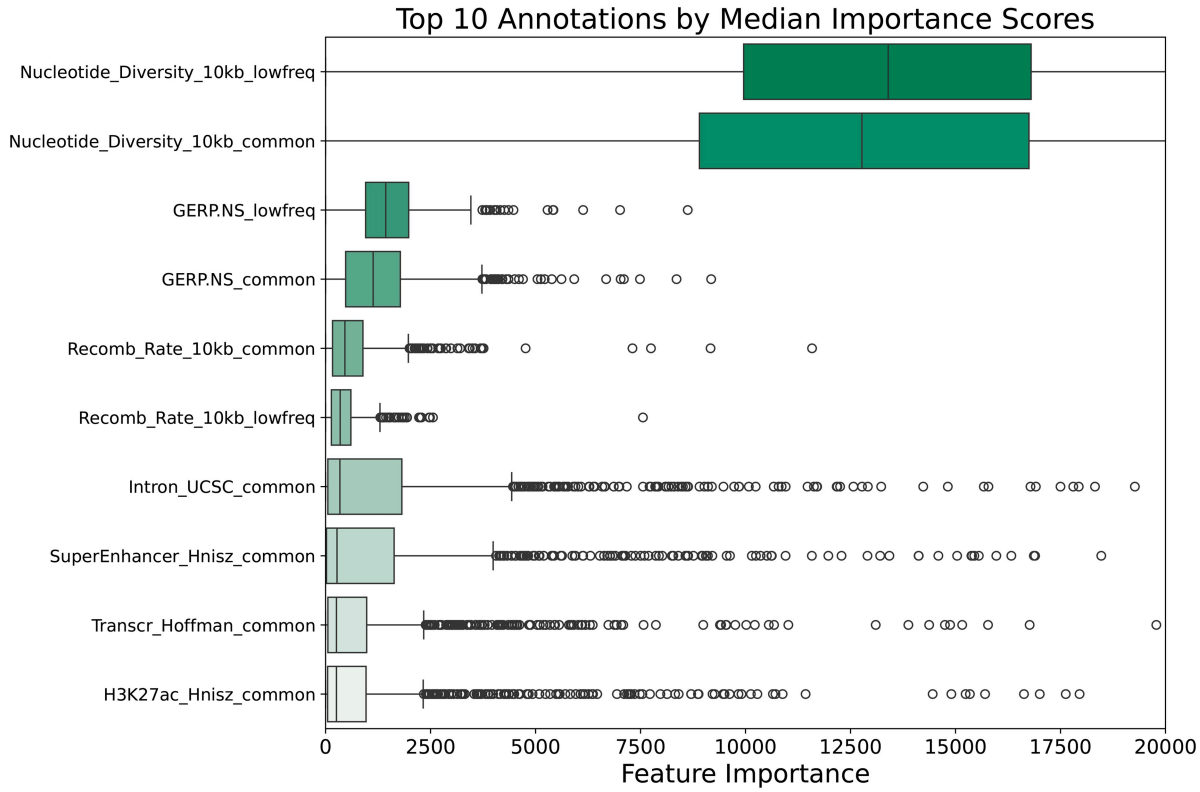

Figure 6. Box plot for Funmap annotation importance scores across 864 genomic regions of four lipid traits (190–374 regions per trait).

(H3K27ac), suggesting the regulatory and epigenetic information are utilized by Funmap to enhance fine-mapping. Interestingly, 6 out of 21 continuous annotations were ranked among the top 10 most relevant annotations, suggesting continuous annotations may contain more informative signals compared to binary annotations. This highlights Funmap's ability to integrating continuous annotations in practical applications.

4 Discussion and conclusion

In this article, we introduced a novel and computationally efficient fine-mapping method, Funmap, to adaptively integrate functional information from a vast amount of SNP annotations. It not only boosts the statistical power of fine-mapping, but also produces well-controlled FDR when considering a large number of functional annotations. Funmap's efficient algorithm allows it to simultaneously identify multiple causal signals while handling hundreds of annotations. Through comprehensive simulations, we showed that Funmap achieved greater statistical power and higher resolution while producing better-controlled FDR. We applied Funmap to identify causal SNPs for 4 lipid traits by integrating 187 functional annotations, yielding substantial power gains. Importantly, we showed that Funmap effectively avoided spurious results, identifying causal SNPs that could be better reproduced.

Our Funmap framework can be extended along several directions. First, while our main analysis focuses on functional features related to genic region, allele frequency, and LD, publicly available resources at single-cell resolution are rapidly growing, serving as new data resources to annotate risk SNPs. Integration of fine-mapping with single-cell genomics, epigenomics, and proteomics data will be a promising direction to reveal the cellular contexts of causal signals (Yu *et al.* 2022). Second, Funmap selects relevant annotations by adopting an ad-hoc ranking procedure, which does not guarantee the control of FDR or type-I error rate when inferring associated annotations. To enable statistically rigorous inference on annotation weights, Funmap's random effects model can be extended by introducing a sparse probabilistic structure, such as the spike-slab model (Ming *et al.* 2018).

Supplementary data

Supplementary data are available at *Bioinformatics* online.

Conflict of interest: None declared.

Funding

Y.Z. was supported by National Natural Science Foundation of China (No. 12301383), Shenzhen Science and Technology Program (No. RCBS20221008093336086), Hetao Shenzhen-Hong Kong Science and Technology Innovation Cooperation Zone Project (No. HZQSWS-KCCYB-2024016) and Longgang District Special Funds for Science and Technology Innovation (No. LGKCSDPT2023002). M.C. was supported by City University of Hong Kong Startup Grant (No. 7200746) and Strategic Research Grant (No. 21300423). J.M. was supported by the National Natural Science Foundation of China (No. 12201219) and Shanghai Key Program of Computational Biology (No. 23JS1400500, 23JS1400800). J.X. was supported by National Natural Science Foundation of China (No. 12401384) and Shenzhen Science and Technology Program (No. RCBS20231211090613024). This research has been

conducted using the UK Biobank Resource under Application Number 96744.

Data availability

Selected gene regions in simulation are available at https://github.com/ZikunY/CARMA/tree/master/Simulation%20Study. Summary data of Lipid-related traits are available at https://nealelab.github.io/UKBB_ldsc/index.html. LD files for UKBB British-ancestry are available at https://nealelab.github.io/UKBB_ldsc/index.html. Functional annotations for real data are available at https://alkesgroup.broadinstitute.org/LDSCORE/baselineLD_v2.1_annots. Summary statistics from GLGC can be downloaded at https://csg.sph.umich.edu/willer/public/glgc-lipids2021.

References

Benner C, Spencer CC, Havulinna AS *et al.* Finemap: efficient variable selection using summary data from genome-wide association studies. *Bioinformatics* 2016;**32**:1493–501.

Bouchard G. Efficient bounds for the softmax function and applications to approximate inference in hybrid models. In: *NIPS 2007* Workshop for Approximate Bayesian Inference in Continuous/Hybrid Systems, Vancouver, Vol. 6. 2007.

Bycroft C, Freeman C, Petkova D *et al.* The UK biobank resource with deep phenotyping and genomic data. *Nature* 2018;**562**:203–9.

Cai M, Xiao J, Zhang S *et al.* A unified framework for cross-population trait prediction by leveraging the genetic correlation of polygenic traits. *Am J Hum Genet* 2021;**108**:632–55.

Carbonetto P, Stephens M. Scalable variational inference for Bayesian variable selection in regression, and its accuracy in genetic association studies. *Bayesian Anal* 2012;**7**:73–108.

Chen W, Larrabee BR, Ovsyannikova IG *et al.* Fine mapping causal variants with an approximate Bayesian method using marginal test statistics. *Genetics* 2015;**200**:719–36.

Consortium EP *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;**489**:57.

Fachal L, Aschard H, Beesley J, *et al.*; ABCTB Investigators. Fine-mapping of 150 breast cancer risk regions identifies 191 likely target genes. *Nat Genet* 2020;**52**:56–73.

Gazal S, Finucane HK, Furlotte NA *et al.* Linkage disequilibrium–dependent architecture of human complex traits shows action of negative selection. *Nat Genet* 2017;**49**:1421–7.

Graham SE, Clarke SL, Wu K-HH, *et al.*; Global Lipids Genetics Consortium. The power of genetic diversity in genome-wide association studies of lipids. *Nature* 2021;**600**:675–9.

Guan Y, Stephens M. Bayesian variable selection regression for genome-wide association studies and other large-scale problems. *Ann Appl Stat* 2011;**5**:1780–815.

Gusev A, Lee SH, Trynka G, SWE-SCZ Consortium *et al.* Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. *Am J Hum Genet* 2014;**95**:535–52.

Hormozdiari F, Kostem E, Kang EY *et al.* Identifying causal variants at loci with multiple signals of association. In: *Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics, California*, 2014, 610–11.

Kichaev G, Yang W-Y, Lindstrom S *et al.* Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLoS Genet* 2014;**10**:e1004722.

Kichaev G, Roytman M, Johnson R *et al.* Improved methods for multi-trait fine mapping of pleiotropic risk loci. *Bioinformatics* 2017;**33**:248–55.

Kundaje A, Meuleman W, Ernst J, *et al.*; Roadmap Epigenomics Consortium. Integrative analysis of 111 reference human epigenomes. *Nature* 2015;**518**:317–30.

Lee Y, Luca F, Pique-Regi R *et al.* Bayesian multi-SNP genetic association analysis: control of FDR and use of summary statistics. BioRxiv, 2018;316471. https://doi.org/10.1101/316471

Ming J, Dai M, Cai M *et al.* LSMM: a statistical approach to integrating functional annotations with genome-wide association studies. *Bioinformatics* 2018;**34**:2788–96.

Schaid DJ, Chen W, Larson NB. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat Rev Genet* 2018;**19**:491–504.

Slatkin M. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* 2008;**9**:477–85.

Strembitska A, Labouèbe G, Picard A *et al.* Lipid biosynthesis enzyme Agpat5 in AgRP-neurons is required for insulin-induced hypoglycemia sensing and glucagon secretion. *Nat Commun* 2022;**13**:5761.

Uffelmann E, Huang QQ, Munung NS *et al.* Genome-wide association studies. *Nat Rev Methods Primers* 2021;**1**:59.

Wang G, Sarkar A, Carbonetto P *et al.* A simple new approach to variable selection in regression, with application to genetic fine mapping. *J R Stat Soc Series B Stat Methodol* 2020;**82**:1273–300.

Weissbrod O, Hormozdiari F, Benner C *et al.* Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat Genet* 2020;**52**:1355–63.

Wen X, Lee Y, Luca F *et al.* Efficient integrative multi-SNP association analysis via deterministic approximation of posteriors. *Am J Hum Genet* 2016;**98**:1114–29.

Yang Z, Wang C, Liu L *et al.* CARMA is a new Bayesian model for fine-mapping in genome-wide association meta-analyses. *Nat Genet* 2023;**55**:1057–65.

Yu F, Cato LD, Weng C *et al.* Variant to function mapping at single-cell resolution through network propagation. *Nat Biotechnol* 2022;**40**:1644–53.

Zhang W, Najafabadi H, Li Y. SparsePro: an efficient fine-mapping method integrating summary statistics and functional annotations. *PLoS Genet* 2023;**19**:e1011104.

Zou Y, Carbonetto P, Wang G *et al.* Fine-mapping from summary data with the "sum of single effects" model. *PLoS Genet* 2022;**18**:e1010299.

Bioinformatics, 2025, 41, 1–12
https://doi.org/10.1093/bioinformatics/btaf017
Original Paper